Trading Volume and Dispersion of Signals [*]

Nikhil Vidhani [†]

June 29, 2020

[*] This is a first draft. I would like to thank my thesis chair Prof. Sankarshan Basu, committee members Prof. Srinivasan Rangan, and Prof. Abhinav Anand and seminar participants at IIM Bangalore for helpful comments.

[†]Indian Institute of Management, Bangalore. *email*: nikhil.vidhani16@iimb.ac.in. Ph: (+91)-797-555-7296

Trading Volume and Dispersion of Signals

Abstract

I propose a new measure of investor disagreement using trading signals from a broad set of return predicting anomalies constructed utilizing the firm's fundamentals and market price. Disagreement is higher for small, growth, and riskier stocks, which exhibit high fundamental uncertainty. The dispersion of trading signals significantly explains the next period trading volume over and above the factors identified in earlier literature. A move from 25th to 75th disagreement percentile predicts 12.7% higher turnover next period. The positive and significant relationship is robust to different specifications, alternative measures of turnover and disagreement, across size groups, and periods. Univariate and bivariate portfolio sorts offer similar evidence. Firm-specific information in the form of management disclosures, analyst recommendations, and media coverage also affects the disagreement-volume relationship. Investors rely less on return anomalies for firms with several sources of value relevant information and vice versa. I find the disagreement return relationship to be stronger for small firms, value stocks, levered firms, and firms with thin analyst following.

Keywords: Disagreement, Return Anomalies, Trading Volume

1 Introduction

Disagreement among market participants is a fundamental motivation to trade. If two traders hold opposing views regarding the future value of an asset, then profit motive would drive them to trade with each other. Disagreement can arise if investors interpret the information contained in stock price and fundamentals differently by using different models of asset valuations. Representing different valuation models by return predicting anomalies documented in the literature, I propose a new method to capture investor disagreement as the dispersion of trading signals generated by these anomalies. An investor following a particular strategy may initiate a buy (sell) trade if the strategy suggests higher (lower) returns in the future. To the extent that investors rely on these return anomalies for their investment decisions, dispersion in trading signals emanating from these anomalies would capture the disagreement among investors regarding asset's future returns. Using a wide range of anomalies constructed from the firm's fundamentals and market price, I find that disagreement significantly explains the next period trading volume over and above the factors identified in prior literature on determinants of the trading volume. To the best of my knowledge, no study has attempted to study the impact of trading signals generated by return anomalies on volume.

Standard rational expectations (RE) asset pricing models with common priors and homogenous interpretation of information (like CAPM) can't generate trading once the hedging and insurance needs of traders are satisfied (Varian (1989)). In a fully revealing RE equilibrium, i.e., where prices reflect all available information, investors without private information are not willing to trade since they anticipate the superior nature of information possessed by other investors (Milgrom and Stokey (1982)). Asymmetric information dries up the liquidity, and the market fails in a manner similar to Akerlof (1978).

To overcome the impossibility of trade in a fully revealing RE setting, researchers assume some exogenous source of randomness either through noise trading (Kyle (1985)), endowment shocks (Diamond and Verrecchia (1981)) or liquidity shocks. Some form of exogenous noise must be assumed, even with differential information, as a means to conceal trading intentions of insiders holding superior information. This assumption makes equilibrium only partially revealing as opposed to fully revealing[1]. For instance, in Kyle (1985), privately informed investors trade due to

[1] In a partially revealing RE equilibrium, agents' private information is only partly reflected in the price. When there are two sources of randomness, i.e., asset's final payoff, and exogenous noise, then prices can't fully reveal the information pertinent to asset's payoff (Diamond and Verrecchia (1981)). Without exogenous noise, asset's payoff related information is fully reflected in price and in that sense prices are fully revealing.

camouflage provided by noise traders.

The assumption of exogenous noise, which creates room for trading, is also an inherent weakness of partially revealing RE equilibrium models. Any interesting patterns in trade emerge as a direct consequence of the assumed structure of noise, limiting the role of differential information in predicting trading patterns (Kruger (2019)). Disagreement models can overcome the shortcomings of RE models. They assume some form of investor belief heterogeneity, which creates disagreement regarding the asset payoff. Agents agree to disagree in equilibrium and hence interpret information differently, giving rise to trade as investors revise their beliefs (Banerjee and Kremer (2010)). Investors can disagree when one group of investors (specialists) acquire value relevant information earlier than the other investors causing them to disagree and hence trade more. This is evident with trading volume spiking at information events like earnings announcements. Disagreement can also arise when investors hold heterogeneous priors or interpret public information like earnings announcements differently. Hong and Stein (2007) give a brief overview of this literature.

Disagreement comes in many flavors. Investors can disagree owing to different access to information, different tastes and endowments, and divergent beliefs. Varian (1989) shows that a mere difference in information, taste, or endowments, without any belief heterogeneity, can't give rise to trading. Investor must have differing beliefs for trading to appear. Investor's beliefs can differ if investors hold different opinions (priors) or they interpret the common information differently. Investor's posterior belief is a function of her prior belief and likelihood function. The latter is akin to how information is interpreted.

The beliefs of market participants are unobservable; hence, disagreement can't be measured directly. For equity markets, brokerage firm analysts, an important albeit limited group of non-participating agents, periodically issue forecasts of earnings and other accounting fundamentals. The dispersion in their forecasts has been extensively used as a proxy for disagreement[2]. Although standard, the analyst forecast based measure suffers from several weaknesses. First, these analysts are sell-side agents employed by brokerage firms who earn commissions on the number of shares sold following the issue of their forecasts. There is also evidence that analyst forecasts are positively biased (Beyer and Guttman (2011)) and exhibit herding behavior. Secondly, this measure is not a market-based measure in that it only includes views of a small class of agents

[2] Kandel and Pearson (1995) isolate flips and divergences of the same analyst's successive forecast surrounding an earnings announcement and attribute their frequency of occurrence to the extent of differential interpretation. More recently, Fischer et al. (2019) regress analysts' forecast on prior forecasts of other analysts and use the extent of deviation of the slope coefficient from one as a measure of disagreement.

(Atiase and Bamber (1994)). Moreover, analyst following is highly skewed towards larger firms and forecast dispersion with small analyst following is imprecise[3]. Finally, the most common forecast issued is for the earnings, and it may differ from beliefs about the asset's future value. Thus forecast dispersion may not accurately capture the disagreement relating asset's future price.

Empirical asset pricing studies of cross-sectional determinants of trading volume are relatively rare as compared to studies of stock returns. We know a great deal about what factors determine returns, but not so much can be said about trading volume[4]. Statistically, trading volume is very different from stock returns. Trading volume is an absolute measure; hence its distribution is highly skewed, exhibits significant autocorrelations and heavy right tails (Ajinkya and Jain (1989)). Moreover, trading volume is on a secular rise ever since, mostly due to the reduced transaction costs[5], increased retail participation, and ease of trading via electronic media. Lo and Wang (2010) tried a host of factors to remove trends and jumps from volume but couldn't get a stationary time-series.

Chordia et al. (2006) is perhaps the first comprehensive empirical study of the cross-sectional determinants of trading activity. They consider several variables capturing liquidity trading, informed trading, and trading due to belief divergence in a single predictive regression framework. More recently, Jacobs and Hillert (2015), building on the empirical model of Chordia et al. (2006), add several other variables related to stock's visibility like S&P membership and advertising expense. They find that firm names appearing in the upper half of the alphabet experience significantly more trading than those in the lower half. Lo and Wang (2010) perform an exploratory investigation of trading volume and suggest other variables like alpha and residual standard deviation from CAPM regression and return auto-correlation to affect trading volume.

In this study, I revisit the disagreement volume relationship using a novel market-based measure of disagreement. Disagreement captured by the dispersion of buy/sell trading signals predicted by several return anomalies identified in prior literature is significantly related to future volume after controlling for previously studied determinants. Measuring disagreement

[3]The mass of analysts following is also skewed towards big firms. On average, the bottom 30% of firms are followed by 2.26 analysts, while 15.32 analysts follow the top 30% of firms. Similarly, whether analysts cover a firm or not is also correlated with firm size. Bottom 30% of firms have an average of 25% coverage, while the top 30% has 55% coverage. Of the firms covered by analysts, 25% of observations have only one or two analysts issuing forecasts.

[4]Lo and Wang (2010) begin their article by highlighting the remarkable amount of studies exploring the field of asset prices while there being no parallel *asset quantities* literature.

[5]Transaction cost is defined as the net loss of simultaneously buying and selling the desired quantity of shares. It includes brokerage fees, spread, and price impact (if any). The average spread per share is a non-decreasing function of the quantity traded since the desired quantity may not be available at the best bid (or ask) price.

by categorizing a wide range of factors' prediction as buy or sell signals reduces the chances of data snooping. To the extent that investors employ trading strategies originating from a variety of return anomalies for valuing the asset's final payoff, more dispersion across signals should be related to increased trading volume.

Using a large panel of US stocks, I find the disagreement measure to be strongly related to trading volume. A one standard deviation increase in disagreement causes roughly 7.7% additional trading volume in the next month. The relationship is robust to different regression specifications, different measures of turnover and disagreement, across time periods, and size deciles. Disagreement-volume relationship is more robust for smaller, riskier and value firms. Univariate and bivariate portfolio sorts offer similar evidence.

2 Review of Literature

Karpoff (1986) classifies theories about how trading volume arises into three classes: volume related to transaction costs (including the bid-ask spread); volume and price changes, and information-driven volume. Volume has been shown to be decreasing in transaction costs (Epps (1976); Cohen et al. (1979)). In general, most prior research shows that volume is positively related to price changes, with volume being higher under price upticks than under downticks (Epps (1976); Copeland (1976)). Pfleiderer (1984) shows that price changes are uncorrelated with trading by speculators with private information but positively related to trading by liquidity-motivated investors. My focus is on information and its effect on volume. Therefore, I provide a more detailed review of this subset of volume-related theories.

Karpoff (1986) presents a model of trade that assumes random pairing where each owner (seller) and nonowner (buyer) are blindly paired. Trade happens when the buyer's demand price exceeds that of the paired seller due to idiosyncratic demand price revisions. Volume increases if agents, in addition to their idiosyncratic revisions, also have independent differential interpretations of information. The differential interpretation further adds to the variance of demand price revisions resulting in increased trading.

Varian (1989) finds that new information causes trade but only through its differential interpretation. When agents interpret information differently, they also interpret the adjusted price (which contains all new information) differently, and hence trading ensues. In a heterogeneous agent model with different priors and interpretation of information signals, Varian finds that the equilibrium trade for an agent is proportional to the extent by which her opinions and interpretations differ from the average opinion and average interpretations respectively.

Kim and Verrecchia (1991) consider a rational expectations setting where there are two rounds of trading with a public disclosure (like an earnings announcement) in between. Investors receive public and private information signals. The precision of private signals differs across agents. After the first round of trade, investors hold pareto optimal portfolios and hence there is no need for a trade absent the public information disclosure. They find that trading volume in the second round is proportional to the absolute price change between two trading periods as well as the differential precision of pre-disclosure privately held information. Following disclosure, investors update their beliefs in different ways. Investors holding superior information place higher weight on their own information and a smaller weight on public disclosure. Informationally disadvantaged investors do the opposite. Differential beleif updation then gives rise to trading. However, the public announcement is interpreted uniformly. Trading arises solely due to differences in precision of preannouncement information across investors.

Harris and Raviv (1993) consider two groups of agents disagreeing on how they interpret commonly available information signals like announcements of dividends, earnings, stock issues, and macroeconomic events using different likelihood functions. Both groups agree on the "direction" of the common information, i.e., whether it is good or bad but differ in the impact the information has on asset's future payoffs. A positive signal makes the responsive group more optimistic, and this group owns all the shares, while a negative signal makes the unresponsive group relatively optimistic resulting in it owning all the shares. Such a mechanism emerges as a result of risk-neutrality and short-sale constraints, where at each period, only one group ends up owning all the shares.

Kandel and Pearson (1995) (hereafter KP) model differential interpretation of public information signals as arising from investors differing in the mean and precision of their likelihood function, i.e., the model to interpret the public signal. KP model the information signal as consisting of two parts: asset payoff and an additive likelihood function, which is interpreted differently across agents. Agents have different means and precisions of the likelihood term. KP find that public announcements (like EA) can have an abnormal trading reaction even when there is no price reaction and attribute this to different likelihoods used by agents to gauge the future asset payoff. In a two-period heterogeneous agent model, KP find that the equilibrium quantity traded is driven by two factors, one independent of price change and proportional to the extent by which agents disagree about the mean of public signal, and the second factor proportional to price change and differences in precision of the prior beliefs and the public signal. Using analysts' revision of earnings forecasts, authors find significant proportions of revisions to be flips or divergences, which can only appear if analysts interpret information differently. Overall, KP provide substantial empirical evidence that investors interpret public signals (EA) differentially evident from the observed patterns in stock trading volumes.

Kim and Verrecchia (1997) include both preannouncement private information as well as event period information. The former is similar to Kim and Verrecchia (1991) while the later represents new private information to be used only in conjunction with public announcement. The private event period information is the differential interpretation of disclosure which may happen if investors differ in their ability to read or find mistakes in the disclosure. Modelling differential interpretation as different likelihood functions is indistinguishable from event period information. A key result is that differential event period information precision can cause trading without accompanying price change. On the other hand, differential predisclosure information precision increases volume in proportion of absolute price change.

Banerjee and Kremer (2010) consider a difference in opinion (DO) model of trading volume where investors disagree about the interpretation of public information. In a DO model, agents agree to disagree in equilibrium and hence interpret information differently, giving rise to trade as investors revise their beliefs. Disagreement can lead to two types of trading volume response, 1) convergence trade when investors had a prior disagreement but agreed on the current information, and 2) idiosyncratic trade when investors agree on prior information but disagree on the current information. A period of high disagreement leads to a higher volume, which falls slowly and hence exhibits clustering.

This paper is most closely related to KP. I study trading signals generated from widely used return predictors, including fundamental ratios, accounting anomalies, and return momentum acting as different models of asset valuation used by investors. To draw an analogy with the model of KP, the public signal can be thought of as the current market price and the different likelihood functions are the different return predicting signals investors use. For instance, one trader may base her trades using the current book value to calculate the market to book ratio while another may use past price to base her trade on return momentum.

On the empirical side, several papers in the accounting literature have linked disagreement with the abnormal trading volume observed during an earnings announcement (EA). However, the focus in that literature is to understand the information content of earnings reports by studying the price and volume reactions surrounding an EA. The information content of the announcement is gleaned through the magnitude of the reaction. Beaver (1968) emphasizes that a price reaction conveys average market-wide changes in expectations, while a volume reaction indicates changes in individual expectations. The change in individual expectations might arise due to differential precision of prior information or increased disagreement following an EA (see Bamber et al. (2010) for a review). Bamber et al. (1997) find incremental explanatory power using three measures of disagreement in explaining abnormal trading volume around EA: dispersion in prior beliefs, differential interpretation of information signal, and consensus effect arising from post information belief dispersion. All three measures are constructed using analyst's forecasts, and in particular, the second measure for differential interpretation is the dispersion in forecast revisions by the same analyst.

Chordia et al. (2006) perform a comprehensive empirical analysis of the cross-sectional determinants of trading activity. They consider several variables believed to impact trading volume in a single predictive regression framework. Trading volume can be associated with several determinants like liquidity trading, informed trading, and trading due to dispersion in beliefs. Liquidity trading arises due to portfolio rebalancing needs and is positively related to the magnitude of past returns. Trading is supposedly higher in more visible stocks where visibility can be proxied by firm size, firm age, price level, and market to book ratio, all of which predict higher trading volume. If the estimation of a firm's fundamentals is uncertain, then this should lead to more learning-based trade as investors correct their estimation errors. Fundamental uncertainty is proxied by absolute earnings surprise, earnings volatility, and CAPM beta. Lastly, analyst forecast dispersion proxies for belief dispersion and the number of analyst following a firm represents the mass of informed agents, both of which positively predicts trading volume.

Carlin et al. (2014) study how disagreement affects asset returns, volatility, and trading volume using the data from the MBS market. Using a VAR framework of disagreement, volatility, and volume, authors find that increased disagreement leads to increased volatility and volume. Volume also increases with higher volatility, but only when disagreement is high. Investors learn from their trades - higher disagreement leads to higher volume, and subsequently, disagreement falls, resulting in a mean-reverting time series for disagreement.

Jacobs and Hillert (2015) find that firm names appearing in the upper half of the alphabet experience significantly more trading than those in the lower half. They build on the empirical model of Chordia et al. (2006) and propose several other variables that might influence trading like advertising expense, 52-week high/low events, idiosyncratic volatility, market model alpha, media coverage, S&P 500, and DJIA membership.

3 Research Design

I study the relation between disagreement and trading volume in a broader setting than just an earnings announcement (EA). Financial markets are continuously bombarded with new information, be it public announcements, dividend announcements, new issues and buybacks, mergers and acquisitions, macroeconomic news, and geopolitical developments. Studying the relationship surrounding an EA is restrictive, given the rich set of the information environment in which firms operate. A comprehensive and parsimonious model of trading activity should explain the determinants of volume at all times.

Limited attention and costly acquisition of information make it difficult for investors to study the entirety of fundamental anomalies. As such investors are likely to anchor their trading decisions on a small set of anomalies. Dispersion in information signals would trigger trading activity since investors using different anomalies as their model of asset valuation would come up with differing estimates. Some investors would interpret the asset to be undervalued while others would find it to be overvalued.

The factor search literature is vast, with over 300 factors identified in Harvey et al. (2016). To the best of my knowledge, no study has combined both the model of trading volume with return anomaly research and attempted to study the impact of these signals on trading volume. To the extent that investors differ in their model to predict future stock prices by way of focusing on a particular signal, then dispersion across the signals as measured by their standard deviation should relate to trading volume. I hypothesize that the relationship is positive and a higher dispersion causes increased trading volume in the future. We expect the coefficient, β_1 to be positive in Rquation 1:

$$Volume_{i,t+1} = \beta_0 + \beta_1 \cdot Signal\_Deviation_{i,t} + \gamma \cdot Controls_{i,t} + u_{i,t+1} \tag{1}$$

The disagreement-volume relationship hinges on the trading reaction following differential interpretation of public information. Smaller firms are less visible, meagerly followed, issue infrequent disclosures, and have an opaque information environment. The amount of public information in the form of management disclosures, analyst recommendations, and media coverage is limited for small firms. Many times the only disclosures are the ones mandated by regulators, which comprise quarterly earnings and annual balance sheet announcements. On the contrary, large firms routinely issue management forecasts, voluntary disclosures, and major sales announcements in addition to earnings and balance sheet disclosures. This leaves stock price and accounting information as the only source of information for small firms. Since not much is known about smaller firms, investors are more likely to estimate future stock prices using widespread anomalies and fundamental ratios. Hence disagreement-volume relationship should be stronger for such firms, and we should expect the coefficient on β_1 to be comparatively higher for small firms.

Investment using accounting fundamentals and related ratios is popular among value investors. Several books on value investing give prominence to price to book ratio as an indicator of value firms[6]. Value investment attempts to estimate the intrinsic value of a stock using information from the balance sheet and profit and loss statements. The intrinsic value is then compared to the current stock price, and a buy (sell) trade is initiated when the intrinsic value is smaller than the current price. Price to book ratio is also a proxy for visibility where firm with high valuations, i.e., growth firms, are often talked about in media and followed more by analysts[7]. Thus the information environment of value firms is limited and hence the use of anomalies would be higher. We should expect the disagreement-volume relation to be stronger for the high book to market firms since these firms fall into the category of value firms and are more likely to be evaluated using return anomalies originating from accounting fundamentals.

Firms with high leverage are riskier for two reasons. First, high debt levels increase the chances of default.

Second, equity owners in highly levered firms are more willing to take risky bets since the high payoff state rewards the stock owners while low payoff state hurts the debt owners. This gives rise to increased agency costs. Investors would be more vigilant of a risky firm, and their trading would be more sensitive to disagreement arising out of the differential interpretation of information. Hence, the disagreement-volume relationship is likely to be stronger for a risky firm.

[6] Lower P/B ratio (or high book to market) signifies that a firm's market price doesn't accurately reflect it's book value and hence is undervalued. Graham (2006) is a popular investing book following value investing tenets.

[7] Number of analyst following and book to market ratio has a rank correlation of -0.18.

3.1 Measuring Volume

Trading volume has seen explosive growth in the last few decades. The monthly cross-sectional average dollar volume (shares traded) has skyrocketed from $3 Million (0.7 Mn) to $1.62 Billion (35.43 Mn) in the period between 1962 to 2019. In the same period, turnover (defined as a ratio of dollar volume to market capitalization) has gone from 19% to 211%[8]. Thus on average, a company's entire shares changed hands more than twice in December 2019. Figure 1 shows a time series of three measures of volume, averaged across all firms every month. The logarithmized version of the series shows a clear time trend in all three measures.

[Insert Figure 1 here.]

There are several ways to measure volume like share volume, dollar volume, or turnover. Section 2 of Lo and Wang (2010) gives a brief exposition of different approaches. Of the three measures, dollar volume depends on the size of the firm; share volume depends on the share price, and turnover is the only measure independent of share price and firm size. Moreover, dollar volume and share volume vary a lot across the sample. The ratio of 75th percentile to 25th percentile observation of dollar volume, share volume, and turnover is 114.7, 59.2, and 6.22 respectively across all NYSE stocks over the 1962-2019 period. Lo and Wang (2010) also find that if the two fund separation holds, then turnover is the most natural measure for studying the relation between trading volume and equilibrium market models like CAPM.

However, turnover is highly non-stationary. As can be seen in Figure 1, log turnover has a clear linear time trend, which means turnover has grown exponentially over the years. To account for the time trend, I include year dummies in all regressions. Chordia et al. (2006) adjust turnover and other non-stationary time-series using an adjustment procedure proposed in Gallant et al. (1992) (GRT). In their study, GRT remove linear and quadratic time trends, monthly calendar dummies, day of the week effects, and trading gaps (due to holidays) from both the mean and variance of the trading volume. The adjustment procedure involves regressing time series of firm wise turnover on the above effects and subsequently regressing squared residuals from the last regression again on the same effects. This removes any trend or calendar effects from both the mean and variance of the turnover time series. In the regression specifications, I do not consider the above adjustments for the following reasons[9]:

[8] The average turnover hit a maximum of 338% in October 2008.

[9] For completeness I include GRT adjusted turnover in one of the regression specifications (Table 6).

- GRT's motivation for using quadratic detrending is mostly based on visual examination of volume series, which exhibited a downward trend in post-depression years followed by an upward trend since the world war II. This gives a nonlinear visual appearance to the volume series. The sample period for my study is from 1962 to 2019, and there is no apparent second-order non-linearity in turnover to consider quadratic detrending.

- In a bid to achieve stationarity, GRT adjustment also removes the effects mentioned above from the variance of turnover, which is an overkill and reduces the analysis to mostly statistical exercise. After two-tier detrending, it is impossible to assign any economic meaning to the "GRT adjusted" turnover.

- Removing calendar effects and seasonalities reduces the efficacy of the model in explaining these interesting patterns. For instance, increased trading in January is believed to originate from tax-induced trading. Removing monthly seasonality in a model studying tax effects on trading activity will destroy any explanatory power.

3.2 Anomalies

In a first, I use the trading signals generated by a host of return predicting anomalies to constitute a measure of disagreement among market investors. Specifically, when different anomalies generate opposing signals regarding the future return from a stock, investors disagree about the future value of the asset. Disagreement arises since investors interpret the available information present in stock price and fundamentals differentially by employing different return anomalies. I use 31 anomalies studied in a recent paper by Linnainmaa and Roberts (2018) and complement that with five momentum anomalies from McLean and Pontiff (2016). Construction of these anomalies only requires data on stock price and annual reports.

Table 1 contains the list of anomalies used in constructing the deviation measure. The 36 anomalies come from seven categories: profitability (6), earnings quality (3), valuation (5), momentum (5), investment (9), financing (6), and distress (2). Column 3 of the table depicts the predicted return association in the future. For instance, with respect to Gross Profitability anomaly, higher gross profit predicts higher returns in the future. Within each category, anomalies generally have a similar sign of predicted relationship. All six profitability anomalies are positively associated with future returns, while 8 out of 9 investment anomalies predict a negative future return. All anomalies predicting a negative relationship as per their original study are scaled by minus one to make the entire set of 36 anomalies to be positively related to future returns.

[Insert Table 1 here.]

Figure 2 gives a correlation heat map for the 36 anomaly signals. The positive correlation is represented by blue circles while negative correlation is shown by red circles. The size of the circle is proportional to the magnitude of the correlation coefficient. A heat map helps visualize the correlation between signals better than a correlation matrix. All categories of signals, except momentum (15-19) are positively correlated within their respective groups[10]. Profitability (1-6) signals have a strong negative correlation with investment (20-28) signals, a slightly negative association with momentum signals, and a positive correlation with distress (35-36) signals. Earnings quality (7-9) signals are positively correlated with investment signals. Valuation (10-14) and investment signals negatively correlate with distress signals.

[Insert Figure 2 here.]

3.3 Measuring Disagreement

Disagreement measure tries to capture the differential interpretation of public information like stock price and accounting fundamentals, where each investor employs a different return anomaly to estimate stock's future performance. As a first step, each of the 36 return anomalies is used to cross-sectionally rank stocks based on their expected future performance. For instance, a stock with a low market to book value is expected to earn higher than average returns in the future; hence for each period, stocks will be reverse sorted on the market to book value. Stocks are divided into three categories such that the top 30% stocks get a buy (+1) signals, bottom 30% stocks get a sell signal (-1), and the rest are in hold category (0). This process is repeated for all 36 return anomalies giving a series of 36 buy/sell/hold signals for each stock in each period. For every firm-month, I measure disagreement as the standard deviation of these signals.

One off-the-shelf measure which fits in the above structure is an index based on the work of Kaplan and Zingales (1997). This "KZ index" has already been adapted for use in large-sample empirical work by Lamont et al. (2001), so we can follow their methodology. By taking this approach, as opposed to building our own measure of equity dependence from scratch, we hope to minimize any concerns about data mining.

[10] Clustering of blue circles along the diagonal indicates this

There are several ways to capture dispersion other than the standard deviation. Absolute deviation uses absolute differences from the mean value instead of squared differences in standard deviation. One limitation of using a broad buy/sell categorization is that other measures of deviation like interquartile range (IQR) and absolute median deviation are not properly defined. To safeguard against the choice of deviation measure and the splitting criterion for assigning buy/sell signals, I check for robustness using absolute deviation as the measure of disagreement as well as using an 80/20 split where top 20% stocks are assigned a buy signal, and bottom 20% get a sell signal.

Let $A_{f,t,s}$ denote the s^{th} anomaly for firm f at time t. $L_{t,s}$ and $H_{t,s}$ are the low and high cross-sectional cut-offs for determining whether the stock gets a sell signal or a buy signal. For the 70/30 split, $L_{t,s}$ ($H_{t,s}$) is the 30$^{\text{th}}$ (70$^{\text{th}}$) percentile of anomaly s across all firms at time t. The corresponding trading signal $T_{f,t,s}$ is given by,

$$T_{f,t,s} = \begin{cases} -1, & A_{f,t,s} < L_{t,s} \\ 0, & L_{t,s} \leq A_{f,t,s} \leq H_{t,s} \\ 1, & A_{f,t,s} > H_{t,s} \end{cases} \tag{2}$$

The mean trading signal, $\overline{T_{f,t}} = \frac{\sum_s T_{f,t,s}}{|S|}$ and the deviation among the signals $\sigma_{f,t} = \frac{\sum_s (T_{f,t,s} - \overline{T_{f,t}})^2}{|T_{f,t}|-1}$, where $|T_{f,t}|$ is the number of anomaly signals for firm f at time t and $|S|$ is the number of anomalies. $\sigma_{f,t}$ is the measure of disagreement and the primary explanatory variable of interest in this study.

In the rest of the paper, I work with 70/30 cross-sectional stock splits[11]. For each period and firm pair, disagreement is the standard deviation of the 36 signals. Figure 3 gives a time-series of average disagreement[12] and its confidence interval (of two standard errors) over time. For each month, averaging is done across all firms. The average signal dispersion over the entire sample period is 0.755, and the standard deviation is 0.103. Unlike turnover, there is neither a trend nor a spike around 2008 in the time-series of disagreement time-series[13]. The stationary appearance of disagreement gives us confidence that the volume-disagreement relationship isn't affected by a common trend[14].

[11] I use 80/20 stock split as a robustness check.

[12] Cross-sectional average of the firm by firm disagreement is $\left(\frac{\sum_f \sigma_{f,t}}{|F_t|}\right)$ where $|F_t|$ is the number of firms at time t.

[13] The unit-root (or the AR(1) coefficient) is 0.92 which is statistically different from 1 thus making the time-series of average disagreement stationary.

[14] Banerjee and Kremer (2010) and Carlin et al. (2014) argue that investors learn from disagreement and periods of high disagreement are followed by periods of low disagreement. Thus disagreement should form a mean-reverting time-series. Figure 3 shows that disagreement varies on a time scale of business cycles. The average annual disagreement

[Insert Figure 3 here.]

3.4 Nature of Disagreement

The ingredients of the disagreement measure are the return anomaly signals based on accounting data and stock price. Disagreement arises when signals diverge about the predicted return for the stock. In this section, I explore the nature of disagreement and how it is characterized and evaluate questions like: Does disagreement vary by the number of buy and sell trading signals? Do all anomalies contribute equally to the disagreement measure? Which industries cluster more when disagreement is low or high? And lastly, how do firm characteristics like size and return vary with disagreement?

Figure 4 presents average disagreement as a function of the difference in the number of buy and sell signals. Disagreement is highest when we have an equal number of buy and sell signals. This is expected since a buy signal is given a value of +1 while a sell signal is given -1. If we think of signals as a sequence of ± 1 Bernoulli trials, then its variance[15] is maximized when the likelihood of buy signals is the same as that of sell signals.

[Insert Figure 4 here.]

Disagreement and Anomalies

The correlation heat map (Figure 2), although helpful in the visual distinction of groups of anomaly signals, doesn't provide any information of how different anomalies affect disagreement. Figure 5 plots the semi-annual average disagreement constructed using all but one group of anomaly signals. If excluding a particular group reduces the disagreement substantially, then that group contributes significantly to the disagreement measure. Removing profitability or earnings quality signals doesn't have any significant change in disagreement. The exclusion of any other groups of anomalies dampens the disagreement with the momentum group affecting it the most. The reduction in disagreement due to skipping one category of anomalies reduces in the last decade.

has an AR(2) representation of $d_t = 0.30 + 1.05d_{t-1} - 0.46d_{t-2} + u_t$ with an R^2 of 0.62. The AR(2) model was chosen based on the AIC criterion. The cycle length of this model (as described in Tsay (2010) pg. 42) turns out to be 9.16 years.

[15] With 70/30 splits, the average number of buy and sell signals across firms is fixed and equal to $0.3 * 36 = 10.8$. However, within a time-period number of buy and sell signals will vary across firms. Since the signals are not independent, the number of buy signals is not a binomial random variable. Hence, finding the analytical variance of the number of buy signals is much more involved and depends on the correlation structure of signals.

Surprisingly, the exclusion of just two distress anomalies reduces average disagreement from 0.752 to 0.728. This is because the two anomalies, O_Score and Z_Score are mostly negatively correlated with other anomalies and hence contribute significantly to the disagreement measure. This is evident from the last two rows of the heat map in Figure 2 where the two distress anomaly mostly have a negative correlation with other anomalies.

[Insert Figure 5 here.]

Disagreement and Firm Characteristics

Next, I look at how firm characteristics vary with disagreement deciles. Many of the firm characteristics are also the controls in regression analysis in section 5. I look at stock returns, size, earnings surprise, forecast dispersion, and other variables across ten disagreement deciles computed at each month. To facilitate comparison across deciles and characteristics, I use cross-sectional ranks of firm characteristics[16]. Figure 6 presents the variation in ranks of firm characteristics over ten disagreement deciles. Subplots 1,2, and 3 contain firm characteristics, subplot 4 has different disagreement measures, and subplot 5 has different turnover measures.

[Insert Figure 6 here.]

Excess return and Fama and French (2015) (FF5) alpha are stable for the first seven deciles and then start falling. The fall is rather sharp for the last two disagreement deciles where FF5 alpha's rank falls by 25 percentiles. Risk measures like CAPM beta and return volatility rise with disagreement while dividends fall sharply.

Fundamental uncertainty captured by earnings surprise and earnings volatility also associates positively with disagreement. Moving from first to last disagreement decile, earnings surprise, and volatility rank change from 38% to 82%. The signals making up the disagreement measure are mostly constructed using fundamental information from the balance sheet and P/L statements. A more uncertain income stream and business forecast cast doubt on the future asset value, and the signals which measure future value will indicate this uncertainty as increased disagreement.

On the other hand, the disagreement measure is largely unrelated to analyst forecast dispersion and the number of analysts following a firm. The mass of analyst following represents the

[16]Stocks are ranked each month by each characteristic and then scaled to fall between 0 and 1. Thus they are equivalent to percentiles. Details of variable construction appear in Appendix A.2.

interest of informed agents in a particular stock and hence is a measure of the level of informed trading in the stock. Trading due to disagreement arises purely from non-informational[17] reasons. The relationship between the number of analysts and disagreement is slightly negative. Forecast dispersion is also a measure of disagreement but is only faintly positively related to fundamental disagreement. Dispersion in analyst forecast represents the dispersion in the private information content of analysts (Barron et al. (1998)), and hence it is a proxy for differential content of private information possessed by analysts. It is thus different in form and structure from the fundamental disagreement, which originates as a result of differential interpretation of commonly available return predicting signals. Across ten disagreement deciles, forecast dispersion rank changes from 30% to 40%. It is important to note that forecast dispersion and fundamental disagreement are two measures of disagreement that differ by their origins and hence capture different types of heterogeneity among investor's trading motives. We should not expect these to be substitutes in a regression setup where disagreement explains some property of asset prices like returns, volatility, or trading activity.

Disagreement is also significantly higher among small and younger firms. Across the ten disagreement deciles, the ranking for firm size and age falls by more than 50 percentiles. O_Score, a measure of distress risk, is also significantly higher for high disagreement firms. Disagreement also associates negatively with book to market ratio where growth firms (having low BTM ratio) tend to be firms with high disagreement. Growth firms derive a major chunk of their value from real options and growth opportunities which makes their revenue projections difficult to estimate. This may lead to higher disagreement among investors about the future asset value of these firms.

Different measures of disagreement move sharply with fundamental disagreement deciles. The particular choice of disagreement should then be irrelevant in explaining trading volume. This is expected since different measures only differ in their mathematical formulae, not in the trading signals generated by return anomalies. Lastly, subplot 5 shows several turnover measures as a function of disagreement deciles. All turnover measures increase gently over the ten disagreement deciles. Turnover adjusted for time trends and calendar dummies ($TURN\_GRT$) increases more steeply. Interestingly, illiquidity captured by Amihud (2002) measure increases sharply with disagreement. Moving across ten disagreement deciles, changes illiquidity by 50 percentiles.

[17]The superior information of the analyst is the private information available only to the analyst utilizing her stock/industry research, management interviews, and information of the macro-economy. This information is different from the information contained in fundamental signals considered in this study. These signals are common information, and investors are assumed to interpret the common information differently, leading to fundamental disagreement.

Kruger (2019) explores the relationship between disagreement and liquidity in detail.

Overall, the average firm in high disagreement decile pays no dividend, experience low excess returns with increased volatility, is small and young, has high earnings uncertainty, has low BTM ratio, experience high distress risk, and illiquidity when compared to a firm in low disagreement decile.

Disagreement and Industry Characteristics

Evidence from figure 6 suggests that disagreement is related to fundamental uncertainty, size, and risk characteristics of firms. Since different industries vary in their risk characteristics, should disagreement be different for firms in different industries? For instance, should disagreement be higher for firms in established industries like utilities that have a constant source of revenues or should it be higher for firms in the pharmaceutical industry whose value depends a lot on drug trials, regulatory approvals, and R&D investments? Looking at which industries cluster at low and high disagreement deciles gives more insight into the nature of disagreement.

Figure 7 gives for each disagreement decile the percentage of firm-month observations belonging to a particular industry. The relative concentration of each industry changes across disagreement deciles. The 48 industries (Fama and French (1997)) are presented in 8 subplots of 6 industries each. Industries are sorted on their relative concentration, and then the smallest group is presented in subplot 1 and the largest in subplot 8. This arrangement puts industries with several firms in the later subplots. Looking at the last four subplots, industries like computers, construction, investment, clothing, electronic chip making, wholesale, drugs and medical equipment, oil, and IT are the ones whose concentration increases with disagreement decile. While industries like electrical equipment, food, office supplies, building materials, machinery, chemicals, and utilities show the opposite trend, i.e., their concentration falls with disagreement deciles.

[Insert Figure 7 here.]

Industries with real options like IT, computers, pharmaceuticals, and oil[18] are harder to value since their future income streams is dependent on the exercise of these options. Similarly, other

[18] Firms in the oil and drug industry derive their real options from exploration activities. Firms in the oil industry have an option to explore a potential oil field, whether to develop the site and finally, when to start drilling and extract petroleum products. Besides, the oil industry is also heavily dependent on global crude prices, which adds an extra degree of uncertainty. The drug development process is also full of real options at various stages, trials, approvals, and patents.

growth industries like financial trading, electronic equipment, and semiconductor chip-making are also highly concentrated in high disagreement deciles. Established industries with a smooth revenue stream like building materials, chemicals, machinery, business supplies (paper), and utilities are easier to value. These industries are concentrated in lower disagreement deciles. Banking and insurance industries seem to have an inverted U-shaped relationship with disagreement in that the maximum concentration occurs at moderate disagreement.

Overall, industries that possess fundamental uncertainty, be it uncertain future cash flows, growth opportunities, patents, global risks, real options, and regulatory hurdles, concentrate at high disagreement. Increased fundamental uncertainty makes valuation difficult, which is reflected in higher disagreement originating from return anomalies.

4 Data and Sample

The first step to construct the disagreement measure is to construct the anomalies. I employ 31 anomalies from Linnainmaa and Roberts (2018) and five momentum anomalies from McLean and Pontiff (2016). I obtain monthly stock returns, share outstanding and volume data from the Centre for Research in Security Prices (CRSP) database. Stock's annual fundamentals are obtained from COMPUSTAT's annual files, and earnings data is acquired from quarterly files. Both CRSP and COMPUSTAT data are accessed from Wharton Research Data Services (WRDS) account. Analyst forecast summary data is obtained from the Institutional Brokers' Estimate System (IBES) database via the Thomson Reuters account. Matching across the datasets is done using CUSIP identifier. Annual accounting data is available from Januaury 1962, quarterly data from July 1971, monthly stock return from December 1925 and analyst forecast data from January 1976. The above availability dictates our sample to start from 1976. The sample ended in 2019. For quarterly earnings data, it is required that an announcement is made within 180 days after the fiscal quarter-end. For calculating analyst forecast dispersion, I require at least two annual forecasts. If forecast dispersion could not be calculated due to the unavailability of data, then previous estimates are carried forward. All these requirements leave us with roughly 400 thousand firm-month observations. Finally, all variables are winsorized at 0.5% on either extreme before estimating a regression.

4.1 Data Snooping

Data snooping is a significant concern in empirical asset pricing research. I try to be as agnostic and comprehensive as I can be regarding my choice of factors. However, certain limitations on data availability restrict the choice of factors. The most comprehensive recent list of fundamental factors is McLean and Pontiff (2016), which employs 97 past determinants of returns. Many of these factors use IBES and other proprietary data, which prohibits their inclusion in my analysis. Computational and time constraints at my end also restrict the universe of return anomalies I can consider while calculating the disagreement measure. Finally, I settle with all 31 factors studied by Linnainmaa and Roberts (2018) and the five momentum anomalies studied in McLean and Pontiff (2016).

The construction of disagreement measures can still be infected with data-snooping concerns. Lo and MacKinlay (1990) argue that portfolios formed using previously studied return characteristics are more prone to data snooping. I offer several justifications as to why the disagreement measure may not be susceptible to data mining:

- I am not aware of any previous work using the dispersion among return predicting anomalies to predict trading volume, and hence the arguments of Lo and MacKinlay (1990) don't apply. Moreover, my work is focused on explaining volume, whereas data snooping as such is attributed widely to return anomalies literature where factor searching is intense (Harvey et al. (2016)).

- I am using only a rudimentary dispersion measure – standard deviation – and as such, this dispersion measure is unlikely to be significantly affected by the inclusion/exclusion of a few factors. Earlier, we have seen that average disagreement depends on the aggregate sum of individual signals correlation, and hence a particular signal can't significantly affect the average level of disagreement.

- In constructing the disagreement measure the exact degree by which a particular signal predicts return is not used. For instance, if prior research has documented that factor A predicts 1% higher returns while factor B predicts only 0.1% higher returns, a stock falling in the "buy" categories for both the factors will get a +1 signal w.r.t. each factor. In other words, looking just at the signal, it is impossible to tell how much of a future return is predicted by the underlying factor. A more involved approach exploiting the degree of return prediction would assign a higher number to signal A than to signal B. This prevents basing the disagreement measure on previously observed returns, precisely the concern raised by Lo and MacKinlay (1990).

- I am choosing a reasonable number of factors in its entirety from a previously published study (Linnainmaa and Roberts (2018)) and complementing it with the full set of momentum-based factors from McLean and Pontiff (2016). Hence, there is no cherry-picking of factors that would fit the data.

- The study covers all firms over the entire duration for which stock's price and fundamental data is available over the CRSP and COMPUSTAT databases, thus nulling the p-hacking argument due to sample selection (Harvey et al. (2016)). The large sample at hand also avoids any false/lucky discovery due to small sample usage.

- I use the multi testing threshold of $|t| \geq 2.78$ for ascertaining significance instead of the usual 1.96 (Harvey et al. (2016)). Moreover, all standard errors are double clustered by year-month and firm[19], and are more conservative.

5 Results and Discussion

I use portfolio sorts and linear regression to establish the relation between disagreement and volume. Since the goal is to explain the determinants of volume, all the right-hand side variables are lagged by one month. Turnover as measured by monthly dollar volume scaled by market capitalization, L_TURN is regressed on the disagreement measure, STD_DEV, and a host of control variables and fixed effect dummies. I use the same set of firm characteristics previously studied in Chordia et al. (2006) as control variables. To control for the time trend in volume and the possibility of it varying across industry types, I include dummy variables for each year of the sample and the 48 industry types identified in Fama and French (1997)[20]. I estimate the below regression (without intercept),

[19] Clustering by time is equivalent to Fama-MacBeth (FM) procedure where time series mean and standard error of per period cross-sectional regression estimates are reported as coefficient estimates and its standard error. In time clustering entire cross-section of a period is taken as one observation, and it preserves the cross-correlation structure of asset returns. Similarly, firm clustering takes entire time-series of a firm as one observation and therefore takes into account the auto-correlation structure of persistent variables like trading volume. This is equivalent to Newey-West (NW) adjustment. In general, double clustering performs better than FM regression combined with NW adjustment as long as there is a sufficient number of clusters (~ 50) in each dimension.

[20] The list of industry codes is available at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_48_ind_port.html

$$L\_TURN_{i,t} = \sum_k \alpha_k \cdot Controls_{k,i,t-1} + \beta \cdot STD\_DEV_{i,t-1} + \sum_d \gamma_d \cdot Dummies_{i,t-1} + \epsilon_{i,t} \quad (3)$$

The controls used are past returns separated into positive and negative returns (RET^+ and RET^-), firm leverage measured by debt to equity ratio (LEV), beta coefficient from firm-level CAPM regression ($CAPM\_BETA$), book to market ratio (BTM), current market price (L_PRC), firm's age as measured by logarithm of number of months since the first trading (L_FAGE), firm size (L_ME), absolute surprise in a firm's earnings and the volatility of earnings ($ESURP$ and $EVOL$), number of analyst following a firm and their forecast dispersion ($NUMEST$ and $FDISP$). A preceding $L\_$ ahead of a variable means its logarithm is used in the regression. A complete description of variable definitions and their construction appears in Appendix A.2.

Table 2 gives the correlations among several control variables and the disagreement measure, i.e., all variables that appear on the right-hand side equation 3. High correlation between RET^+ and RET^- is inconsequential since both of them are never non-zero simultaneously. L_ME is strongly correlated with several other variables. Its correlation with L_PRC is mechanical; with $NUMEST$ is due to correlation between analyst following and firm size. $ESURP$ and $EVOL$ have stock price in the denominator and that might be driving the negative correlation with L_ME. To avoid multicollinearity concerns, I drop L_ME from the regression design. The other significant correlation is between $ESURP$ and $EVOL$, which may also be driven by their common denominator. Disagreement (STD_DEV) is positively correlated with LEV, $CAPM\_BETA$, $ESURP$, and $EVOL$. It is higher for smaller (L_ME) and younger (L_FAGE) firms. Interestingly, disagreement is much more correlated with negative returns than it is for positive returns. Overall, the evidence from correlations is consistent with findings from Figure 6.

[Insert Table 2 here.]

Table 3 presents summary statistics of the explanatory variables outlined above. Additionally, I also present summaries for several other possible measures of disagreement and volume. The derived turnover measures like L_TURN_D, L_TURN_ILLIQ, VW_L_TURN, and EQ_L_TURN have high standard deviation despite their mean being close to 0. This is expected since all these measures are the residuals of regressing turnover on some other measure. Variables representing counts like NUM_FLIPS, NUM_DIV, and L_FAGE, truncated variables like RET^+ and RET^-, and earnings-related variables $ESURP$ and $EVOL$ have high skewness and kurtosis. Except for L_FAGE and RET^-, all are positively skewed. Most of the turnover measures and STD_DEV are relatively symmetric. Pure turnover measures like L_TURN, L_TURN_1d, and L_TURN_5d are slightly platykurtic (kurtosis < 3) while derived measures like L_TURN_D, EW_L_TURN, and L_TURN_GRT are slightly leptokurtic. Analyst forecast variables ($NUMEST$ and $FDISP$), disagreement rank (STD_DEV_R), and detrended turnover (L_TURN_D) have very small excess kurtosis.

[Insert Table 3 here.]

In Table 4, I regress[21] log turnover (L_TURN) on several combinations of controls and the disagreement measure (STD_DEV). The first specification is the base regression from Chordia et al. (2006)[22]. The sign and relative importance of coefficients in the specification (1) are in harmony with Chordia et al. (2006) results. Except for RET^- and L_FAGE, all variables have a positive coefficient. The effect of BTM and $EVOL$ on turnover is not statistically significant. All regressions have industry and year fixed effects, and the standard errors are double clustered by firm and month. Petersen (2009) finds that if there are at least 50 clusters in each dimension, then double clustering of standard errors is closer to true standard errors than Fama-MacBeth regressions with Newey West correction. Since there are many more firms and almost 700 months of observations (1962 to 2019), there will always be enough clusters, even with sub-sample analyses.

Specifications 4-6 include the disagreement measure to the list of controls in specifications 1-3. The quantum of all controls except L_PRC gets reduced after including STD_DEV. L_FAGE loses significance after including disagreement. The coefficient on STD_DEV is stable across specifications, and in the full model (specification (4)), R^2 increases by roughly 3.5% (0.205/0.198). I report two R^2 for each regression: adjusted R^2 is for the full model including fixed effects, while Within R^2 is for within-group variation with fixed effects projected out. $\%R^2$ explained is the percentage of unexplained variation explained by including disagreement in the regression[23]. Disagreement roughly explains 1% of the unexplained variation of the full model. The coefficient

[21] All regressions reported in this paper are executed on R using the lfe package (https://cran.r-project.org/web/packages/lfe/lfe.pdf) and verified on Stata's reghdfe package (http://scorreia.com/help/reghdfe.html).

[22] Specification (10) in Table 3 of Chordia et al. (2006) is similar to the specification (1) in Table 4.

[23] Let R^2_w and $R^2_{w/o}$ be the R^2 from a regression that includes and excludes disagreement respectively. Then $\%R^2$ explained is the ratio of $R^2_w - R^2_{w/o}$ and $1 - R^2_{w/o}$.

on STD_DEV is 0.736, and its standard deviation is 0.101, which means that a 1 SD increase in STD_DEV predicts the next month's log turnover to increase by 0.074 (0.736*0.101). This is equivalent to 7.72% ($e^{0.074}$) increase in turnover. If I exclude earnings (forecast) related variables, the net effect increases to 7.86% (8.25%) for 1 SD change in disagreement.

[Insert Table 4 here.]

The next set of regressions evaluates the effect of disagreement on more immediate daily and weekly turnover. In specification (4) of Table 4, henceforth called the base regression, next month turnover is predicted using variables constructed at the end of the current month. In Table 5, I explore the relationship between the end of the month disagreement and turnover measured during the first day and the first week[24] of the next month. The six regressions of Table 5 can be grouped into two groups: first three, which doesn't include disagreement and the last three, which includes STD_DEV. Comparisons can be made both across the group (with and without disagreement) and within each group (turnover over daily, weekly, and monthly intervals). The coefficient on STD_DEV and the R^2 gently increases over the regressions 4-6 providing evidence that disagreement affects turnover across different time intervals. The coefficient increase from 0.644 to 0.742, indicating an increasing role of disagreement in explaining volume. Forecast dispersion, on the other hand, is relatively stable across different time horizons. The within model R^2 changes substantially from regression 4 to 5 but also from regression 1 to 2 which means the additional explanatory power primarily comes from variables other than STD_DEV. The coefficients of visibility and liquidity related variables like RET^+, RET^-, L_PRC and $NUMEST$ decrease sharply in magnitude as we move from explaining daily to monthly turnover. This may be due to the effect of liquidity and visibility being more short termed than disagreement and fundamental uncertainty. For instance, portfolio rebalancing following substantial return changes should materialize in the next few days and not spread over the entire month.

[Insert Table 5 here.]

The strong disagreement-volume relationship may be driven by the choice of the dependent or independent variable. To fully convince ourselves about the validity of results, I assess the

[24] I define the first week of the month as the first five trading days immediately following the end of the previous month.

effect of trying different measures of turnover on the same set of explanatory variables and the disagreement measure STD_DEV. In Table 6, I present regression results from seven different turnover measures. Specification (1) is the base regression. Except for specification (2), all others show a positive and significant relation with turnover. The coefficient on STD_DEV vary based on the dependent variable, but the broad consensus is that a higher disagreement of anomaly signals predict a higher turnover in next month. Results of regression (2) are troubling since the coefficient on RET^{+} and RET^{-} are reversed, which is completely unexpected and in stark contrast to all other regressions. Coefficients on other variables are also severely attenuated. This is mostly due to the choice of dependent variable wherein L_TURN_{t-1} is subtracted from L_TURN_{t}. This takes away all interesting variation from L_TURN since it has a highly persistent time-series. It also imposes a coefficient of one on the first autocorrelation parameter, whereas in the sample it is 0.89. In specification (3), the dependent variable is GRT adjusted log turnover, which is the same as used by Chordia et al. (2006). In regression (5), the dependent variable is the residual from regressing L_TURN on Amihud (2002) Illiquidity measure. To substantiate that the channel through which disagreement affects turnover is not related to liquidity, removing any liquidity related variable from turnover should not affect the disagreement-turnover relationship. Lastly, in specifications (6) and (7) the dependent variables are the residuals from regressing L_TURN on value-weighted and equally-weighted log market turnover respectively. This removes any market-wide component in L_TURN. In regressions 2-7, it is difficult to assign any interpretation to the coefficient since the dependent variable is heavily transformed. With the forecast dispersion ($FDISP$), the results are mixed. The coefficient loses significance when the dependent variable is detrended turnover, GRT adjusted turnover, and market adjusted turnover.

[Insert Table 6 here.]

Similarly, in the next table, I evaluate the effect of varying the way disagreement is measured. In Table 7, L_TURN is regressed on controls and different sets of disagreement measures whose definition appears in Appendix A.2. For brevity, only the disagreement variables are shown in the regression table. The coefficients are all positive and significant, except for NUM_FLIPS. Specification (1) is our base regression, and the R^2 for all specifications except (5) are very similar to base regression R^2. All measures except the number of flips and divergences explain roughly 1% of unexplained variation in turnover with absolute deviation (ABS_DEV) explaining the most. A one SD change in STD_DEV, ABS_DEV and PC_DEV predicts 7.2%, 7.7% and 7.9% higher trading volume in next month respectively. A one SD change in LO_CORR_SD and HI_CORR_SD, the two disagreement measures from non-overlapping sets of low and high correlation signals, explains 3.1% and 4.8% of next month's volume. The overall explanatory power is similar to other measures, but highly correlated signals have a relatively higher explanatory power. The evidence in Table 7 provides strong support for the relation between disagreement among anomaly signals and turnover. It doesn't matter how disagreement is measured; the relationship is consistently positive and significant. On the other hand, the number of flips and divergences, as defined in Kandel and Pearson (1995) and applied to anomaly signals rather than analyst forecasts, doesn't provide any significant explanation for turnover.

[Insert Table 7 here.]

My results may be driven by outliers, wherein a small number of observations have a high disagreement as well as large turnover. The standard in empirical literature to account for outliers it to winsorize the variables. At each month, I winsorize the sample based on the independent variables used in the regression at 0.5% on either extreme. To further solidify the results from previous tables, I also perform rank regressions where the respective cross-sectional rank of a given variable is used rather than the variable itself. The details of how ranks are calculated are present in Appendix A.2. Using ranks have some additional advantages: (i) Outliers do not drive the hypothesized relationship, (ii) interpreting the results become more intuitive since both the explanatory and dependent variable are ranks, and, (iii) different variables can be compared with each other and also over time. In Table 8, I present rank regressions where log turnover and its rank is regressed on ranks of controls and the rank of disagreement measure. A suffix of $\_R$ represents the rank of that variable. For comparison, I report regression without STD_DEV_R as well. Comparing specifications (2) and (4) with specifications (1) and (3) respectively shows that the magnitude of almost all control variables declines after including STD_DEV_R. Specification (4) is the most robust regression since all variables are represented as their cross-sectional ranks. Moving from 25th to 75th disagreement percentile predicts next month turnover rank to be higher by 4 percentiles ($0.081 \times (0.75 - 0.25)$). The inclusion of disagreement improves the explantory power of turnover regression by 1.24%. Similarly, from regression (2), moving from 25th to 75th disagreement percentile predicts 12.7% ($e^{0.239*0.5}$) higher turnover in next month.

[Insert Table 8 here.]

Finally, to investigate whether the result is driven by the size of the firm or a particular time period within the entire sample duration, I perform robustness check for size terciles and nine non-overlapping 5-year subperiods[25]. The latter also tests the the sample performance of the hypothesized relation between disagreement and volume. The three size portfolios are made using 70/30 NYSE breakpoints done separately for each month. For sub-period analysis, the sub-periods start in 1975 and continue until 2019, making nine 5-year periods. Analyst forecast data is available from January 1976 and quarterly earnings data from July 1971; thus, sub-periods can only extend till 1975 using 5-year periods. Even then the first sub-period between 1975 and 1979 has a smaller sample size compared to the rest of the sub-periods.

Table 9 presents the regressions for three size terciles. Specifications 1,3 and 5 are without disagreement, while specifications 2,4 and 6 include disagreement. R^2 is highest for small firms and lowest for big firms. The coefficient on STD_DEV also follows a similar trend where it is most substantial for smaller firms. This substantiates the hypothesis that investors in smaller firms tend to use return anomalies more for their investing needs, and hence disagreement arising from these anomalies is reflected strongly in next month's trading volume. Across the three size groups, a one SD change in disagreement predicts next month's turnover to be higher by 13.4%, 12%, and 5.8% for small, medium, and big stocks respectively. For forecast dispersion, similar estimates are 1.6%, 1.6%, and 0.9%. The magnitude of the coefficient on forecast dispersion is smaller by a factor of 3 when comparing to the base regression (specification (4) of Table 4). It appears that the explanatory power of forecast dispersion is driven by its comovement with firm size and deteriorates substantially when the sample is divided into size-based portfolios. Effect of visibility related variables like L_PRC and $NUMEST$ on turnover is much higher for smaller firms.

[Insert Table 9 here.]

Table 10 has nine regressions for the nine sub-periods. Regression (1) has a smaller sample size due to partial availability of forecast data, and hence the results for this sub-period are less reliable. The coefficient on STD_DEV is significant and positive in all sub-periods and ranges from 0.317 to 1.061. A one SD change in disagreement predicts up to 11.2% higher volume next month in the period 2005-2009. STD_DEV explains almost 1.75% of the unexplained variation for

[25] Internet Appendix also contains robustness checks for leverage and book to market terciles.

this period. On the contrary, the coefficient on $FDISP$ is much smaller in magnitude as compared to specification (4) of Table 4.

[Insert Table 10 here.]

6 Portfolio Sorts

Linear regression imposes a linearity assumption on the population regression function. In general, the relation may be nonlinear. Portfolio sorts can further discern any nonlinearities. Apart from univariate and bivariate sorts, higher-order sorts are harder to visualize as well as suffer from the curse of dimensionality. As the number of variables to sort on increases, the total number of portfolios rises exponentially. For instance, using terciles (3 portfolios) for each variable, the base regression with 12 variables will require more than 500 thousand portfolios (3^{12}). Increasing the number of portfolios reduces the number of firms in each portfolio, and subsequently, the inference is weakened. Hence I restrict portfolio sorts to univariate and bivariate only.

Table 11 shows mean turnover and changes in mean turnover over ten deciles sorted by several control variables and STD_DEV. At the beginning of each month, stocks are assigned portfolio deciles based on their cross-sectional rankings. The variables used to form deciles are lagged by one month. Column 1 is the mean L_TURN in the first decile; column 2 is the amount by which L_TURN increases as we move from the first decile to second decile. Similarly, for the rest of the columns. The effect of liquidity related variables like RET^+ and RET^- is consistently strong except that extreme negative returns reduce turnover. The negative coefficients on RET^- in the second row are due to $RET^- \leq 0$ by construction. Of the visibility related variables, L_PRC affects turnover positively, and L_FAGE affects negatively. BTM proxies for both visibility (growth stocks are more visible) and value investing (investors relying heavily on fundamentals). For the smaller BTM deciles, the effect is negative consistent with the visibility interpretation, but for the last decile, turnover increases significantly consistent with the value investing story. The effect of $NUMEST$ and $FDISP$ is much higher for smaller deciles. $CAPM\_BETA$ proxying for systematic risk, also affects volume consistently. Surprisingly, $ESURP$ and $EVOL$, other fundamental uncertainty measures, affect turnover negatively. Finally, disagreement variables mostly positively affect turnover. Although both $FDISP$ and STD_DEV loose significance in some of the deciles.

[Insert Table 11 here.]

When it comes to two-way portfolio sorts, there are two approaches: independent sorts and dependent sorts. For independent sorts, in the first stage, univariate portfolios are formed independently for both sorting variables, and then their intersection gives bivariate portfolios. The order of sorting is irrelevant. However, if the two variables are correlated, then the firms will be disproportionately allocated across portfolios, and inference becomes less reliable. Dependent sorts can solve this problem where one variable is used to form univariate sorts first, and then another variable is used to sort within the portfolios from the first stage. This makes sure that each portfolio has a similar number of firms. The order of sorting is important here. With dependent sorts on variables A and then B, we can study the effect of B while controlling for A. I use dependent sorts for the bivariate portfolio results. Chapter 5 of Bali et al. (2016) discusses portfolio sorts in detail.

In Table 12, I present average L_TURN over 3×3 portfolio terciles sorted first on one of the control variables and then on disagreement measure STD_DEV. For each control variable tercile, I present the average turnover for first disagreement tercile and then the successive tercile increase in turnover as we move to second and thrid disagreement terciles. The three control variable terciles are grouped as three broad columns; within each column, the average turnover and its changes over disagreement terciles are shown. For instance, $T_{i,j} - T_{i,j'}$ is the difference between disagreement terciles j and j' within control variable tercile i.

Out of 66 tercile changes, 46 are positive, and 12 are negative at the 1% significance level giving strong evidence of the relation between turnover and disagreement after individually controlling for various controls. The average tercile change across different rows (controls) is negative only for $CAPM\_\beta$ and highly positive for L_PRC, $NUMEST$, and $FDISP$. Thus controlling for visibility, informed trading, and differential information content does not change the explanatory power of fundamental disagreement for next period turnover. This further supports the view that STD_DEV and $FDISP$ capture two different types of investor disagreement where the former is the differential interpretation of common information like accounting fundamentals, and stock price and the latter represents the different private information possessed by informed traders or specialists. The average tercile change across broad columns is highest for middle terciles ($T_{22} - T_{21}$ and $T_{23} - T_{22}$) where out of 22 tercile changes, only two are negative. Thus, the effect of disagreement is maximum when the control variables are away from either extreme.

[Insert Table 12 here.]

7 Conclusion

I construct a measure of investor disagreement using a broad set of return predicting anomalies as trading signals and find it to predict trading volume positively. Return anomalies enter as likelihood function in investor's pricing model, and differences in likelihood function give rise to dispersion in beliefs about asset's payoff giving rise to increased trading. To avoid data snooping concerns, I pick all anomalies mentioned in Linnainmaa and Roberts (2018) and top it with momentum anomalies from McLean and Pontiff (2016) and create a rudimentary measure of disagreement: standard deviation of buy/sell signals emanating from the return anomalies.

Disagreement is higher for small, growth, and riskier stocks, which exhibit high fundamental uncertainty. After controlling for prior determinants of trading volume, a move from 25th to 75th disagreement percentile predicts 12.7% higher turnover next period. Similar results also hold if instead of monthly turnover, I use the next day or next week's turnover. The positive and significant relationship is robust to different specifications, different measures of turnover, different measures of disagreement (using return anomaly signals), across size groups, over different time periods, and using rank regressions. Univariate and bivariate portfolio sorts offer similar evidence.

This paper contributes to the volume literature in following dimensions: (i) it uses a new, market-based measure of disagreement drawing from the vast literature of return anomalies, (ii) it adds to the relatively scarce literature on empirical determinants of trading volume and complement it with investor disagreement, and, (iii) provides strong empirical evidence of positive and significant relationship between disagreement and turnover robust to several specifications.

References

Abarbanell, Jeffery S, and Brian J Bushee, 1998, Abnormal returns to a fundamental analysis strategy, *Accounting Review* 19–45.

Ajinkya, Bipin B, and Prem C Jain, 1989, The behavior of daily stock market trading volume, *Journal of Accounting and Economics* 11, 331–359.

Akerlof, George A, 1978, The market for "lemons": Quality uncertainty and the market mechanism, in *Uncertainty in Economics*, 235–251 (Elsevier).

Altman, Edward I, 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance* 23, 589–609.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Atiase, Rowland K, and Linda Smith Bamber, 1994, Trading volume reactions to annual accounting earnings announcements: The incremental role of predisclosure information asymmetry, *Journal of Accounting and Economics* 17, 309–329.

Bali, Turan G, Robert F Engle, and Scott Murray, 2016, *Empirical asset pricing: The cross section of stock returns* (John Wiley & Sons).

Bamber, Linda Smith, Orie E Barron, and Douglas E Stevens, 2010, Trading volume around earnings announcements and other financial reports: Theory, research design, empirical evidence, and directions for future research, *Research Design, Empirical Evidence, and Directions for Future Research (February 27, 2010)* .

Bamber, Linda Smith, Orie E Barron, and Thomas L Stober, 1997, Trading volume and different aspects of disagreement coincident with earnings announcements, *Accounting Review* 575–597.

Banerjee, Snehal, and Ilan Kremer, 2010, Disagreement and learning: Dynamic patterns of trade, *The Journal of Finance* 65, 1269–1302.

Barbee Jr, William C, Sandip Mukherji, and Gary A Raines, 1996, Do sales–price and debt–equity explain stock returns better than book–market and firm size?, *Financial Analysts Journal* 52, 56–60.

Barron, Orie E, Oliver Kim, Steve C Lim, and Douglas E Stevens, 1998, Using analysts' forecasts to measure properties of analysts' information environment, *Accounting review* 421–433.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *The Journal of Finance* 32, 663–682.

Beaver, William H, 1968, The information content of annual earnings announcements, *Journal of Accounting Research* 67–92.

Beyer, Anne, and Ilan Guttman, 2011, The effect of trading volume on analysts' forecast bias, *The Accounting Review* 86, 451–481.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *The Journal of Finance* 43, 507–528.

Bradshaw, Mark T, Scott A Richardson, and Richard G Sloan, 2006, The relation between corporate financing activities, analysts' forecasts and stock returns, *Journal of Accounting and Economics* 42, 53–85.

Carlin, Bruce I, Francis A Longstaff, and Kyle Matoba, 2014, Disagreement and asset prices, *Journal of Financial Economics* 114, 226–238.

Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2006, The cross-section of expected trading activity, *The Review of Financial Studies* 20, 709–740.

Cohen, Kalman J, Steven F Maier, Robert A Schwartz, and David K Whitcomb, 1979, Market makers and the market spread: A review of recent literature, *Journal of Financial and Quantitative Analysis* 14, 813–835.

Cooper, Michael J, Huseyin Gulen, and Michael J Schill, 2008, Asset growth and the cross-section of stock returns, *The Journal of Finance* 63, 1609–1651.

Copeland, Thomas E, 1976, A model of asset trading under the assumption of sequential information arrival, *The Journal of Finance* 31, 1149–1168.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *The Journal of Finance* 61, 1605–1643.

De Bondt, Werner FM, and Richard Thaler, 1985, Does the stock market overreact?, *The Journal of Finance* 40, 793–805.

Diamond, Douglas W, and Robert E Verrecchia, 1981, Information aggregation in a noisy rational expectations economy, *Journal of Financial Economics* 9, 221–235.

Dichev, Ilia D, 1998, Is the risk of bankruptcy a systematic risk?, *The Journal of Finance* 53, 1131–1147.

Epps, Thomas W, 1976, The demand for brokers' services: The relation between security trading volume and transaction cost, *The Bell Journal of Economics* 163–194.

Fama, Eugene F, and Kenneth R French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F, and Kenneth R French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fischer, Paul, Chongho Kim, and Frank Zhou, 2019, Disagreement about fundamentals: Measurements and consequences, Technical report, Working paper.

Gallant, A Ronald, Peter E Rossi, and George Tauchen, 1992, Stock prices and volume, *The Review of Financial Studies* 5, 199–242.

Gao, Meng, and Jiekun Huang, 2019, Informing the market: The effect of modern information technologies on information production, *Review of Financial Studies, Forthcoming* .

Graham, Benjamin, 2006, *The Intelligent Investor* (Harper Business).

Harris, Milton, and Artur Raviv, 1993, Differences of opinion make a horse race, *The Review of Financial Studies* 6, 473–506.

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2016, … and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Haugen, Robert A, Nardin L Baker, et al., 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.

Hong, Harrison, and Jeremy C Stein, 2007, Disagreement and the stock market, *Journal of Economic perspectives* 21, 109–128.

Jacobs, Heiko, and Alexander Hillert, 2015, Alphabetic bias, investor recognition, and trading behavior, *Review of Finance* 20, 693–723.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of Finance* 48, 65–91.

Kandel, Eugene, and Neil D Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831–872.

Kaplan, Steven N, and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints?, *The quarterly journal of economics* 112, 169–215.

Karpoff, Jonathan M, 1986, A theory of trading volume, *The Journal of Finance* 41, 1069–1087.

Kim, Oliver, and Robert E Verrecchia, 1991, Trading volume and price reactions to public announcements, *Journal of Accounting Research* 29, 302–321.

Kim, Oliver, and Robert E Verrecchia, 1997, Pre-announcement and event-period private information, *Journal of accounting and economics* 24, 395–419.

Kruger, Samuel, 2019, Disagreement and liquidity, *Available at SSRN 3189188* .

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica: Journal of the Econometric Society* 1315–1335.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1994, Contrarian investment, extrapolation, and risk, *The Journal of Finance* 49, 1541–1578.

Lamont, Owen, Christopher Polk, and Jesús Saaá-Requejo, 2001, Financial constraints and stock returns, *The review of financial studies* 14, 529–554.

Linnainmaa, Juhani T, and Michael R Roberts, 2018, The history of the cross-section of stock returns, *The Review of Financial Studies* 31, 2606–2649.

Lo, Andrew W, and A Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *The Review of Financial Studies* 3, 431–467.

Lo, Andrew W, and Jiang Wang, 2010, Stock market trading volume, in *Handbook of Financial Econometrics: Applications*, 241–342 (Elsevier).

Lockwood, Larry, and Wikrom Prombutr, 2010, Sustainable growth and stock returns, *Journal of Financial Research* 33, 519–538.

Loughran, Tim, and Jay W Wellman, 2011, New evidence on the relation between the enterprise multiple and average stock returns, *Journal of Financial and Quantitative Analysis* 46, 1629–1650.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2007, The new issues puzzle: Testing the investment-based explanation, *The Review of Financial Studies* 21, 2825–2855.

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Milgrom, Paul, and Nancy Stokey, 1982, Information, trade and common knowledge, *Journal of Economic theory* 26, 17–27.

Novy-Marx, Robert, 2012, Is momentum really momentum?, *Journal of Financial Economics* 103, 429–453.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Ohlson, James A, 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 109–131.

Petersen, Mitchell A, 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *The Review of Financial Studies* 22, 435–480.

Pfleiderer, Paul, 1984, *The volume of trade and the variability of prices: A framework for analysis in noisy rational expectations equilibria* (Graduate School of Business, Stanford Univ.).

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *The Journal of Finance* 63, 921–945.

Sloan, Richard G, 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* 289–315.

Soliman, Mark T, 2008, The use of dupont analysis by market participants, *The Accounting Review* 83, 823–853.

Spiess, D Katherine, and John Affleck-Graves, 1999, The long-run performance of stock returns following debt offerings, *Journal of Financial Economics* 54, 45–73.

Thomas, Jacob K, and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163–187.

Titman, Sheridan, KC John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677–700.

Tsay, Ruey S, 2010, *Analysis of financial time series* (John wiley & sons).

Varian, Hal R, 1989, Differences of opinion in financial markets, in *Financial Risk: Theory, Evidence and implications*, 3–37 (Springer).

Xing, Yuhang, 2007, Interpreting the value effect through the q-theory: An empirical investigation, *The Review of Financial Studies* 21, 1767–1795.

A Appendix

A.1 Construction of Anomalies

All anomalies are constructed using monthly CRSP files and annual fundamentals from COMPUSTAT. I also provide the CRSP/COMPUSTAT variable name in the anomaly definitions. All anomalies are computed for each firm for every period (year-month). Anomalies constructed using only the annual fundamental data are repeated 11 times, i.e., they remain the same over a period of 12 months. If the predicted relationship of an anomaly is negative i.e., a higher value of anomaly predicts lower future returns, then I multiply the anomaly by -1 so that the relationship becomes positive. Subscript t represents the current time period. $\Delta x_t \equiv x_t - x_{t-1}$ and $\overline{x_t} \equiv \frac{x_t + x_{t-1}}{2}$. For momentum anomalies, $ret.\{b\}t\{a\} \equiv \prod_{t=-a}^{t=-b}(1+ret_t)-1$, where $a \leq b$ and ret_t is the return in month t. Book value of a firm's equity is defined as stockholder's equity plus deferred taxes minus preferred shares (Fama and French (1992)). $BE_t = seq_t + txditc_t - pstkrv_t$. If seq is not present, then $(ceq + pstk)$ is used. If either of ceq or $pstk$ is not present, then $(at - lt)$ is used. The market value of equity is the product of shares outstanding and share price: $ME_t = (prc_t/cfacpr_t) * (shrout_t * cfacshr_t)$. If a security's return is not available, then its delisting return is used from the CRSP monthly stock events file. Only firms traded on NYSE, AMEX or NASDAQ ($exchd \in 1,2,3$) having a share code ($shrcd$) of 10 or 11 are considered. Missing return and volume data in an otherwise continuous series are filled with zeros. Accounting data for a firm performing its operations in a year y is matched with trading data of June of year $y+1$ and carried forward 11 months, i.e., the same annual fundamental data is used from June of year $y+1$ to May of year $y+2$ (Fama and French (1992)). Below is the definition of 36 anomalies used in this study:

1. **Gross Profitability** is constructed as revenues minus cost of goods sold scaled by total assets. $GrProf = \frac{revt_t - cogs_t}{at_t}$. Table 1 in Novy-Marx (2013) predicts a positive relationship between gross profitability and future returns.

2. **Operating Profitability** is revenues divided by the sum of cost of goods sold, interest expense and selling, general and admin (SGA) expenses. $OpProf = \frac{revt_t}{cogs_t + xint_t + xsga_t}$. Fama and French (2015) Table 1 Panel B finds a positive association between operating profitability and returns.

3. **Return on Assets** is defined as income before extraordinary items divided by total assets.

$RoA = \frac{ib_t}{at_t}$. Table 3 in Haugen et al. (1996) provides evidence of positive relationship between return on assets and returns.

4. **Return on Equity** is income before extraordinary items divided by book value of equity. $RoE = \frac{ib_t}{BE_t}$. Haugen et al. (1996) Table 3 documents a positive relation between RoE and returns.

5. **Profit Margin** is operating income after depreciation scaled by revenues. $PfMg = \frac{oiadp_t}{revt_t}$. Table 4 of Soliman (2008) finds a positive relationship of profit margin with returns.

6. **Change is Asset Turnover** is constructed as annual change in the ratio of revenues to assets. $ChgAssTurn = \Delta(\frac{revt_t}{at_t})$. Soliman (2008) Table 4 gives and evidence of positive relationship between change in asset turnover and returns.

7. **Accrual anomaly** predicts that stocks with lower accruals earn abnormally high returns in future (Table 6 and 7, Sloan (1996)). Accruals are defined as earnings minus the cash components of earnings divided by average assets. $Accr = \frac{(\Delta act_t - \Delta che_t) - (\Delta lct_t - \Delta dlc_t - \Delta txp_t) - dp_t}{at_t}$. Since the predicted relationship is negative, I transform the last equation by taking the negative of RHS as accrual anomaly.

8. **Net Operating Assets** is defined as the difference of operating assets and operating liabilities scaled by lagged total assets. $NOA = \frac{(at_t - che_t) - (at_t - dlc_t - dltt_t - BE_t)}{at_{t-1}}$. Hirshleifer et al. (2004) Table 5 finds a negative relationship of NOA with returns.

9. **Net Working Capital changes** (annual) are negatively associated with future returns as depicted in Table 7 of Soliman (2008). Net working capital is current assets minus current liabilities. $ChgNWC = \frac{\Delta(act_t - che_t) - \Delta(lct_t - dlc_t)}{at_t}$.

10. **Book to Market ratio** is book value equity divided by market value of equity. Table 6 of Fama and French (1992) finds that stocks with high book to market ratio earns higher future returns. $BTM = \frac{BE_t}{ME_t}$.

11. **Cash Flow to Price ratio** as the name suggests is the sum of earnings and depreciation (the cash flow of a firm) divided by its market equity. $CFP = \frac{ib_t + dp_t}{ME_t}$. Lakonishok et al. (1994) Table 4 find a positive relation between future returns and CFP ratio.

12. **Earnings to Price ratio** is positively related to future returns (Table 4, Lakonishok et al. (1994)). $EP_t = \frac{ib_t}{ME_t}$. The anomaly originally appeared in Basu (1977).

13. **Enterprise Multiple** is defined as the ratio of enterprise value to operating cash flow. $EntMult = \frac{ME_t + dlc_t + dltt_t + pstkrv_t - che_t}{oibdp_t}$. Loughran and Wellman (2011) Table 2 depicts a negative relation between enterprise multiple and future returns.

14. **Sales to price ratio** is negatively associated with returns (Table 2, Barbee Jr et al. (1996)). $SP = \frac{revt_t}{ME_t}$.

15. **Short-term momentum** ($ret.6t2$) is the cumulative buy and hold return from $t-6$ to $t-2$ where t is the current month. Jegadeesh and Titman (1993) in Table 7 find a positive association between short-term momentum and future returns.

16. **Lagged Momentum** ($ret.12t7$) is cumulative buy and hold return from $t-12$ to $t-7$ where t is the current month. Table 1 Novy-Marx (2012) gives evidence of positive association with returns.

17. **Short-term reversal** ($ret.1t1$) is just the last month return. Table 7 of Jegadeesh and Titman (1993) finds a negative relation with future returns.

18. **Momentum reversal** ($ret.18t13$) is the annual buy and hold return starting 18 months prior. Table 7 of Jegadeesh and Titman (1993) finds a negative relation with future returns.

19. **Long-term reversal** ($ret.60t13$) is the buy and hold return from $t-60$ to $t-13$. De Bondt and Thaler (1985) document a negative association.

20. **Asset Growth** is defined as relative change in total asset compared to last year. $AssGr = \frac{\Delta at_t}{at_{t-1}}$. Cooper et al. (2008) in Table III present evidence of negative relation between asset growth and future returns.

21. **Inventory Growth** is the ratio of change in inventory scaled by average assets. $ChgInvt = \frac{\Delta invt_t}{\overline{at}_t}$. Thomas and Zhang (2002) in Table 1 document a negative relationship with returns.

22. **Sustainable Growth** is defined as the relative change in book value of equity compared to last year. $SustGr = \frac{\Delta BE_t}{BE_{t-1}}$. Lockwood and Prombutr (2010) Table 6 finds a negative relation between sustainable growth and future returns.

23. **CAPX Growth** is the relative increase in capital expenditure compared to average expenditure two years before. $CapxGr = \frac{capx_t - \overline{capx}_{t-1}}{\overline{capx}_{t-1}}$. Table 2, Panel B in Abarbanell and Bushee (1998) depict a negative association with future returns.

24. **Growth in Sales minus inventory** is the ratio of relative growth in revenues scaled by relative growth in inventories. $SalesGr\_InvtGr = \frac{revt_t - \overline{revt_{t-1}}}{\overline{revt_{t-1}}} / \frac{invt_t - \overline{invt_{t-1}}}{\overline{invt_{t-1}}}$. Abarbanell and Bushee (1998) in Panel B of Table 2 find a positive association with future returns.

25. **Investment Growth** is the relative growth in capital expenditure. $InvstGr = \frac{\Delta capx_t}{capx_{t-1}}$. Xing (2007) Table 4 depicts negative relation with returns.

26. **Abnormal Capital Investment** is the growth in capital expenditure divided by revenues with respect to its previous three-year average. $AbCapInvst = \frac{capx_t / revt_t}{\frac{1}{3} \cdot \sum_{s=t-3}^{s=t-1} capx_s / rev_s}$. Titman et al. (2004) in Table 6, find it to be negatively related to future returns.

27. **Investment to Capital Ratio** is negatively associated with returns (Table 4, Xing (2007)). $IK = \frac{capx_t}{ppent_{t-1}}$.

28. **Investment to Asset Ratio** is negatively associated with returns (Section 4, Lyandres et al. (2007)). $IA = \frac{\Delta ppent_t + \Delta invt_t}{at_{t-1}}$.

29. **Debt Issued Indicator** is one if a firm issues net debt in year and zero otherwise. $DebtIssueInd = 1_{\Delta dlc_t + \Delta dltt_t > 0}$. Spiess and Affleck-Graves (1999) find it to be negatively associated with future returns.

30. **Leverage is defined** as long term debt to book value of equity. $LEV = \frac{dltt_t}{ME_t}$. Bhandari (1988) in Table 2, Panel A find it to be positively related to future returns.

31. **One-year Share issuance** is the increase in number of shares outstanding with respect to last year. $ShIssue\_1 = \frac{shrout_t \cdot cfacshr_t}{shrout_{t-1} \cdot cfacshr_{t-1}}$. Table 3 in Pontiff and Woodgate (2008) depicts a negative relation with future returns.

32. **Five-year Share issuance** is the increase in number of shares outstanding with respect to shares outstanding five years ago. $ShIssue\_5 = \frac{shrout_t \cdot cfacshr_t}{shrout_{t-5} \cdot cfacshr_{t-5}}$. Table 3 in Daniel and Titman (2006) depicts a negative relation with future returns. The anomaly originally appeared in Daniel and Titman (2006).

33. **Total External Finance** is the net financing raised in a year including both equity and debt. Bradshaw et al. (2006) in Table 5 find it to be negatively related to future returns. There are two versions of external finance measure:

    a. $ExtFin = \frac{(ShIssue\_1_t - 1) \cdot ME_t + \Delta dlc_t + \Delta dltt_t - dvc_t}{at_t}$

b. $ExtFin2 = \frac{sstk_t - prstkc_t + dltis_t - dltr_t + dlcch_t - dv_t}{at_t}$

34. **O-Score** is a measure of distress by Ohlson (1980). It is the fitted value of a logistic regression. $O\_Score = -1.32 - 0.407 \cdot \log(\frac{at_t}{cpiind_t}) + 6.03 \cdot \frac{lt_t}{at_t} - 1.43 \cdot \frac{act_t - lct_t}{at_t} + 0.076 \cdot \frac{lct_t}{act_t} - 1.72 \cdot 1_{lt_t - at_t > 0} - 2.37 \cdot \frac{ib_t}{at_t} - 1.83 \cdot \frac{oiadp_t}{lt_t} + 0.285 \cdot 1_{ib_t < 0,\ ib_{t-1} < 0} - 0.521 \cdot \frac{\Delta ib_t}{|ib_t| + |ib_{t-1}|}$. Dichev (1998) in Table 4 find O-score to negatively predict future returns.

35. **Z-score** is a distress measure by Altman (1968). $Z\_Score = 1.2 \cdot \frac{act_t - lct_t}{at_t} + 1.4 \cdot \frac{re_t}{at_t} + 3.3 \cdot \frac{ni_t + xint_t + txp_t}{at_t} + 0.6 \cdot \frac{ME_t}{lt_t} + 1.0 \cdot \frac{revt_t}{at_t}$. Table 3 in Dichev (1998) finds a positive relation with returns.

A.2 Variable Definitions

The number of analysts following a firm ($NUMEST$) and dispersion in analyst forecast ($FDISP$) uses the I/B/E/S data available from Thomson Reuters. I use the EPS summary file for the US companies and restrict the sample to annual forecasts (having $fpi == 1$). To account for missing data due to analyst unfollowing a firm, both $NUMEST$ and $FDISP$ are repeated forward until forecast data is available. Earnings surprise ($ESURP$) and volatility of earnings ($EVOL$) are constructed using quarterly fundamentals from COMPUSTAT. Since earnings are reported once in three months, both $ESURP$ and $EVOL$ are repeated for two months to get a monthly measure.

A preceding $L\_$ before a variable name implies the logarithm of that variable. A trailing $\_D$ implies deterministic linear time detrending of the concerned variable[26]. Suffixing a variable with $\_R$ represents the cross-sectional ranking of that variable[27]. Lastly, a subscript t represents the current period (month), $t-1$ the previous period, and so on. For instance, ME_t is the current market value of equity, L_ME_{t-1} is the natural log of previous month's market equity, ME_D_{t-2} is the detrended market equity lagged by two periods, and $L\_ME\_R_t$ is the cross-sectional rank of current log market equity. Note that detrending and ranking are performed on the logarithmized variable and not the other way round. At times it is convenient to work with quantiles of variables like portfolio sorts and dummy variable regressions. A trailing $\_Dec$ and $\_Ter$ represents deciles and terciles of the concerned variable respectively.[28] Table 13 gives the definitions of variables used in the regressions:

[Insert Table 13 here.]

[26]Detrending is performed by the following regression: $X_t = \alpha + \beta \cdot t + \epsilon_t$. X_D then equals ϵ_t

[27]At each month t the concerned variable is sorted and ranks are assigned. A smaller value gets a smaller rank. To ease comparison of different variables across different time periods, the ranks are scaled (separately for each month) to fall between 0 and 1.

[28]The boundaries for deciles are $10, 20, \ldots, 90$ percentiles while for terciles the boundaries are 30 and 70 percentiles. Both deciles and terciles are computed cross-sectionally for each time period.

IA Internet Appendix

IA.1 Disagreement and Correlations

Disagreement and signal correlations are tightly related to each other. If two signals are perfectly corrleated then they have zero disagreement within them while if they are perfectly negatively correlated then they have maximum deviation. Thus correlation and disagreement are negatively related to each other. To see the connection more generally, assume that for each signal a fraction p of the stocks fall in the buy category, q in sell category and the rest $1-p-q$ in the hold category. $T_{f,t,s}$ can then be modelled as a random variable such that,

$$T_{f,t,s} = \begin{cases} 1, & \text{w.p.} \quad p \\ -1, & \text{w.p.} \quad q \\ 0, & \text{w.p.} \quad 1-p-q \end{cases}$$

We can drop the firm identifier f as all the avergaes and correlation will be across all firms. The signals also have a time-varying correlation structure given by $Corr(T_{ts}, T_{ts'}) = \rho_{tss'}$. Its straightforward to show that, $\mathbb{E}[T_{ts}] = p-q$, $\quad \mathbb{E}[T_{ts}^2] = p+q \quad$ and $\quad \text{var}[T_{ts}^2] = p+q-(p-q)^2$. $\mathbb{E}[T_{ts}T_{ts'}] = \mathbb{E}[T_{ts}]\cdot\mathbb{E}[T_{ts'}] + \text{cov}(T_{ts}, T_{ts'}) = (p-q)^2(1-\rho_{tss'}) + (p+q)\rho_{tss'}$. The mean signal, $\overline{T_t} \equiv \frac{1}{N}\sum_s T_{ts}$, where N is the total number of anomaly signals, has a mean of $\frac{p-q}{N}$ and $\mathbb{E}[\overline{T_t}^2] = \frac{1}{N^2}\sum_s\sum_{s'}\mathbb{E}[T_{ts}T_{ts'}] = \frac{1}{N^2}\sum_s\sum_{s'}\left((p-q)^2(1-\rho_{tss'}) + (p+q)\rho_{tss'}\right) = (p-q)^2 + \frac{C_t}{N^2}\left((p+q)-(p-q)^2\right)$ where $C_t \equiv \sum_s\sum_{s'}\rho_{tss'}$ is sum of all N^2 correlations. The disagreement between signals is defined as, $S_t \equiv \frac{1}{N-1}\sum_s(T_{ts} - \overline{T_t})^2 = \frac{1}{N-1}\left(\sum_s T_{ts}^2 - N\overline{T_t}^2\right)$. Taking expectation we get,

$$\mathbb{E}[S_t] = \frac{N}{N-1}\left(1 - \frac{C_t}{N^2}\right)\left(p+q-(p-q)^2\right)$$

$\mathbb{E}[S_t]$ is the cross sectional average of disagreement at time t, which is maximized when $p = q = 0.5$. However in the time-series it depends only on signal correlations and the dependence is only through sum of all indivudal correlations, C_t. This is reassuring in terms of the construction of disagreement where its dependence on one particular signal is limited to how the signal is correlated with all other signals. Skipping or adding a few signals shouldn't vastly change the disagreement. The dependence of disagreement on correlation also gives a natural demarcation of low and high disagreement regimes. We can partition the entire set of signals ($\mathscr{S}$) into two disjoint subsets $\mathscr{S}^{lo}$ and $\mathscr{S}^{hi}$ such that $C_t^{hi} - C_t^{lo}$ is maximized. $\mathscr{S}^{lo}$ represents a set of signals which have small correlations among themselves while $\mathscr{S}^{hi}$ is the set of highly correlated signals. Since $\frac{\partial \mathbb{E}[S_t]}{\partial C_t} < 0$, disagreement within $\mathscr{S}^{lo}$ should be higher on average. We are interested in examining which set of signals predicts trading volume more strongly. Signals within $\mathscr{S}^{hi}$ are usually of the same sign becuase of high correlation. Investors would expect these signals to move together and any disagreement would come as a surprise. Thus whenver signals in this set disagree, it should cause a heightened trading response because disagreement within $\mathscr{S}^{hi}$ is less likely and when it occurs it creates more uncertainty about the asset's future performance[29].

IA.2 Correlation between disagreement and signals

I explore the correlations between disagreement and signals. According to Figure 5, momentum anomalies contributes most to the disagreement level but it doesn't tell how does signal variation relate to variation in disagreement. If one particular signal or a class of signals is recommending sell then what does it tells us about the disagreement? Will it be higher, lower or unaffected? Figure 8 gives the correlation map of disagreement and signals. Profitability, value, security issuance and distress signals are mostly negatively correlated while earnings quality is positively correlated. Other signals groups doesn't seem to have a clear direction of association with disagreement. Of the 36 signals, Earnings to Price ratio has a correlation of -0.280 and Net Operating Assets has a correlation of 0.158 with disagreement. The average correlation across all 36 signals is -0.028. If we exclude distress anomalies from disagreement this correlation goes upto -0.126. It is an interesting finding that most of the signals correlate negatively with disagreement. The signals are return predicting anomalies and a higher value of signal predicts higher future return[30]. Thus a negative association between anomalies and disagreement hints at a negative disagreement return relationship.

[Insert Figure 8 here.]

[29] Unfortunately, we can't say anything about var$[S_t]$ within the framework of the above model. This requires computation of covariances between signal pairs, i.e. cov$(T_{ts}T_{ts'}, T_{ts''}T_{ts'''})$. Estimating these from data is also not possible since with 36 signals the total number of estimations every period would be ${}^{36}C_4 = 58905$.

[30] All anomalies which predict a negative association with future returns are scaled by -1 so as to make them positively associate with future returns. The predicted association is as per the original source which documented the anomaly. The complete list of anomalies and the predicted relationship is present in Table 1.

IA.3 Robustness Checks

As a robustness check, I perform disagreement-volume regression across portfolios partitoned by book to market and leverage terciles. I also look at NASDAQ stocks and different stock splitting rules. Lastly, I present some more results from univariate portfolio sorts.

IA.3.1 Book to market splits

Tables 14 gives the regression results for BTM terciles. The coefficient on STD_DEV is highest for high BTM stocks i.e. value stocks. Across the three terciles, a one SD change in STD_DEV predicts 4.9%, 8.9% and 12.5% higher turnover in next month respectively. Adj. R^2 and $\%R^2$ explained also increases with BTM terciles. Thus, disagreement arising from fundamental anomalies has more explanatory power for value stocks. This also provides evidence in favour of the hypothesis that investors in value stocks primarily use return anomalies for their trading decisions and hence disagreement among the anomalies strongly predicts next month's trading volume.

[Insert Table 14 here.]

IA.3.2 Leverage splits

Table 15 shows the coefficient on STD_DEV is highest for the largest leverage tercile. A one SD change in disagreement predicts 8.3%, 9.2% and 12.1% higher trading volume next month across the three leverage terciles. In the first tercile, $\%R^2$ explained is only 0.1% while in third tercile it increases to almost 2% indicating that fundamental anomalies affect trading volume much more strongly for high leverage firms. Since leverage proxies for a firm's operational and default risk, a higher coefficient on STD_DEV provides evidence for the increased vigilance and responsiveness of investors in risky stocks.

[Insert Table 15 here.]

IA.3.3 Analyst Following splits

Analysts' recommendations can substitute for anomalies as investment advice. If a firm is not followed by many analysts then investors would rely more on anomalies from accounting fundamentals and stock price for their trading decisions. Increased reliance on anomalies should increase the trading volume reaction arising from disagreement within thise anomalies. Thus we should expect the disagreement-volume relationship to be stronger for firms with small analyst following. Table 16 gives the results for $NUMEST$ terciles. The coefficient on STD_DEV is highest for the first $NUMEST$ tercile. A one SD change in disagreement predicts 10.9%, 6.8% and 6.3% higher trading volume next month across the three $NUMEST$ terciles. Surprisingly the significance of forecast dispersion vanishes in the second and third terciles.

[Insert Table 16 here.]

IA.3.4 Firm Age Splits

I find the disagreement-volume relationship to be stronger for younger firms. Table 17 gives the results for firm age terciles.

[Insert Table 17 here.]

IA.3.5 Report length splits

Complexity of quarterly and annual reports could also pose difficulties in gauging the performance of firms. These disclosures provide additional information (other than B/S and P/L items) regarding the current state of the firm. A lengthier, i.e., wordier, disclosure presents a complex and convoluted snapshot of the firm performance and hence difficult to interpret. For such cases falling back to a simpler anomalies based investment could be favoured. We should expect the use of return anomalies to increase for lengthier disclosures and subsequently the disagreement volume relationship would be strengthened. Table 18 presents results from terciles made on the length of the disclosure document[31]. As expected, the disagreement coefficient is highest for the third tercile. The reduction in forecast dispersion coefficient is also largest for the thrid tercile.

[Insert Table 18 here.]

[31] Disclosure length is for 10-Q and 10-K documents is obtained from a summary file compiled by Bill McDonald. It is available from 1993 onwards at https://sraf.nd.edu/textual-analysis/resources/#LM_10X_Summaries. Quarterly (10-Q) and annual (10-K) entries are merged with COMPUSTAT data using the *cik* identifier.

IA.3.6 EDGAR implementation

Securities and Exchange Commision (SEC) mandated online filing of corporate disclosures through Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system in a phased manner from April 1993 through May 1996[32]. Prior to this rule, mandatory disclosures, like 10-K and 10-Q, were not readily accessible to the general public. Availability of accounting fundamentals present in annual reports is essential for constructing the return anomalies. In lieu of easy access to fundamentals, the use of return anomalies would be limited. Post EDGAR implementation, the use of firm disclosures for investment decisions should increase, subsequently increasing the reliance on return anomalies. With increasing use of return anomalies for trading decisions, disagreement arising from these anomalies will have an increased impact on trading activity. Hence, we should expect the disagreement volume relationship to be stronger post EDGAR implementation.

Figure 9 gives the annual regression coefficients for disagreement. The average coefficient before 1993 is 0.39, while after 1996 it increases to 0.72. The level of coefficients shifts to a higher value around EDGAR implementation. The percent of unexplained variation (unreported) also show similar trend. Table 19 gives regression results from two disjoint periods: pre and post EDGAR implementation. As hypothesized, the coefficient on disagreement is much larger after EDGAR implementation. A one SD change in disagreement predicts 3.9% and 9% higher trading volume next month for pre and post EDGAR periods respectively.

[Insert Figure 9 here.]

[Insert Table 19 here.]

IA.3.7 Shareholding Pattern

The disagreement-volume relationship is accentuated in presence of short-sale constraints. Institutional owners act as supplier for shorting stock and hence low ownership suggests increased short sale constraints. Using another measure ofr short sale constraints - short interest - I find similar results.

Table 20

[Insert Table 20 here.]

[32] Appendix A of Gao and Huang (2019) gives the time line of the staggered EDGAR implementation. Companies were required to file electronically in 10 groups.

IA.3.8 Bid-Ask Terciles

Bid-ask spread can be used to proxy for information symmetry.

IA.3.9 Uncertainty Terciles

Volatility of daily returns captures the firm specific uncertainty. We may also use idiosyncratic volatility instead.

IA.3.10 Time since announcement

Accounting information is disclosed annually while trading happens continously based on the same information. During annual announcements, investors would use firm relevant information disclosed in announcement as well as the stock price and fundamentals to assess the future value of the stock. However, once the announcement becomes stale the only source of new information is the stock price. Since many of the return anomalies (like valuation and momentum) use stock price in conjunction with balance sheet items, the use of return anomalies should be diminished. Conversely, return anomalies should be more important after announcement. Table 21 presents regression results with four subsamples based on time since EDGAR 10-K filing. The disagreement coefficient is smallest when the anouncement is very stale and a new announcement is about to happen. It is largest in seven to nine months post announcement. Over the four regression specifications, a one SD change in disagreement predicts 9.2%, 9.2%, 9.8% and 8.0% higher trading volume next period.

[Insert Table 21 here.]

IA.3.11 NASDAQ stocks

NASDAQ stocks are structurally different from NYSE/AMEX stocks. The exchange was constituted in 1971 with electronic stock market. The stocks at NASDAQ exchange tend to be young and small technology firms. As of December 2018, the average NYSE/AMEX firm is 2.6 times bigger and 9 years older than the average NASDAQ firm. The evidence in Tables 9 and 17 suggest that small and young stocks have a bigger disagreement coefficient. Since NASDAQ stocks generally have both these characteristics, we should expect to see much larger coefficients as compared to base rgeression (specification (4) of Table 4). Table 22 below gives the regression summary for

NASDAQ stocks. Not only is the coefficient on disagreement higher than base regression, the explanatory power is also much higher. Disagreement explains upto 4% of unexplined variation in turnover after controlling for other covariates. A one SD change in disagreement predicts 11.8% higher turnover in the full model (Table 22 specification (4)).

[Insert Table 22 here.]

IA.3.12 Different stock splits

I have performed the analysis in this study using 70/30 stock splits on NYSE/AMEX universe to construct disagreement. To safeguard against the choice of the particular splitting criterion, I perform robustness check using 80/20 and 50/50 splits. Additionally, I also calculate disagreement using all stocks from NYSE, AMEX and NASDAQ exchanges. Table 23 gives the results. The biggest regression coefficient appears when all stocks are used in a 80/20 split to form disagreement where the inclusion of disagreement improves unexplained variation by 2.1%. 50/50 split gives insignificant results with no increase in R^2. The base regression used in this study is the NYSE/AMEX 70/30 split of specification (6). A one SD change in ${}^{ALL}STD\_DEV^{80/20}$, ${}^{ALL}STD\_DEV^{70/30}$, ${}^{NYSE}STD\_DEV^{80/20}$ and ${}^{NYSE}STD\_DEV^{70/30}$ predicts next month's turnover to be higher by 10.8%, 8.7%, 9.7% and 7.8% respectively.

[Insert Table 23 here.]

IA.3.13 Univariate sorts: different disagreement measures

Table 24 gives the changes in average turnover over deciles made on different disagreement measures. Majority of the decile changes are psoitive where out of the total 63 decile changes, 8 are negative and 27 are positive at 1% significance level. Across deciles, the evidence is strong in earlier disagreement deciles and mixed in later deciles.

[Insert Table 24 here.]

IA.3.14 Univariate sorts: different turnover measures

Table 25 present portfolio averages of different turnover measures with portfolios sorted on STD_DEV. Turnover is measured at time t while portfolios are made at time $t-1$. The effect of disagreement is strongest for adjusted turnover measures like L_TURN_GRT and L_TURN_ILLIQ. For other measures, the effect diminishes in later deciles. Overall, the broad majority of decile changes are positive and significant for all turnover measures. There is evidence that illiquidity also rises with disagreement (Figure 6, subplot 5). This can have a negative effect on turnover as disagreement rises.

[Insert Table 25 here.]

Plots and Tables

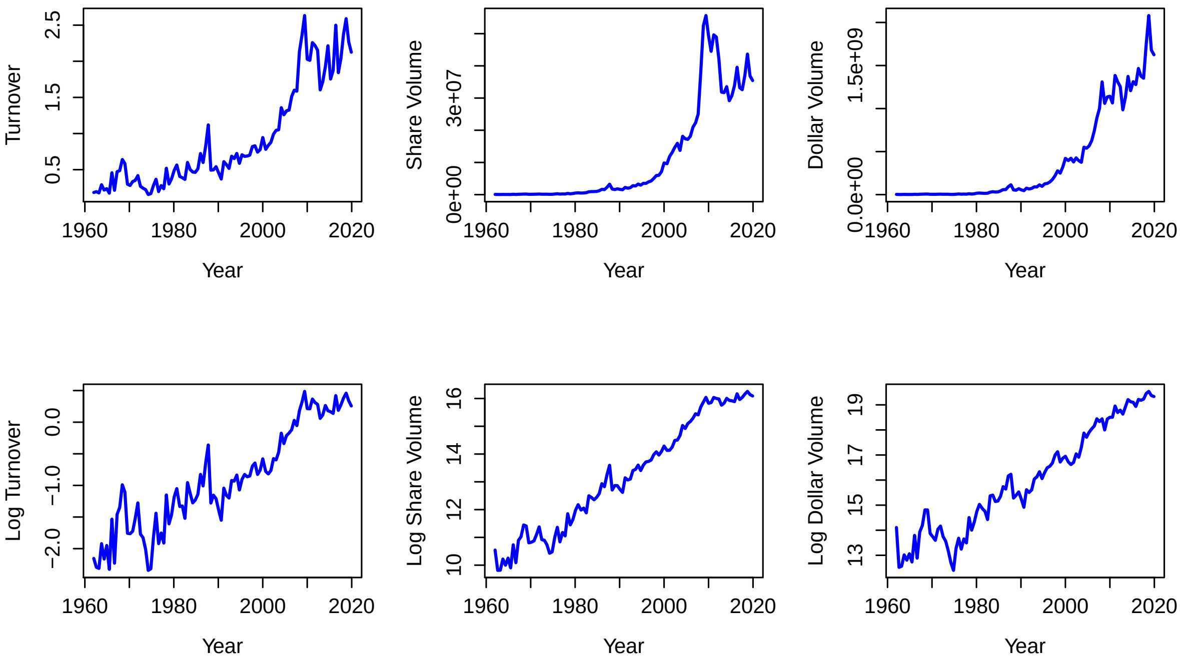

Figure 1: For each month, cross-sectional averages of three trading activity measures across all NYSE stocks are plotted. Averaging is done each month over the period 1962-2019. Trunover is the ratio of dollar volume to dollar market capitalization. Share volume is the number of shares traded. Top three plots use raw variables while bottom plots use logs of corresponding variables.

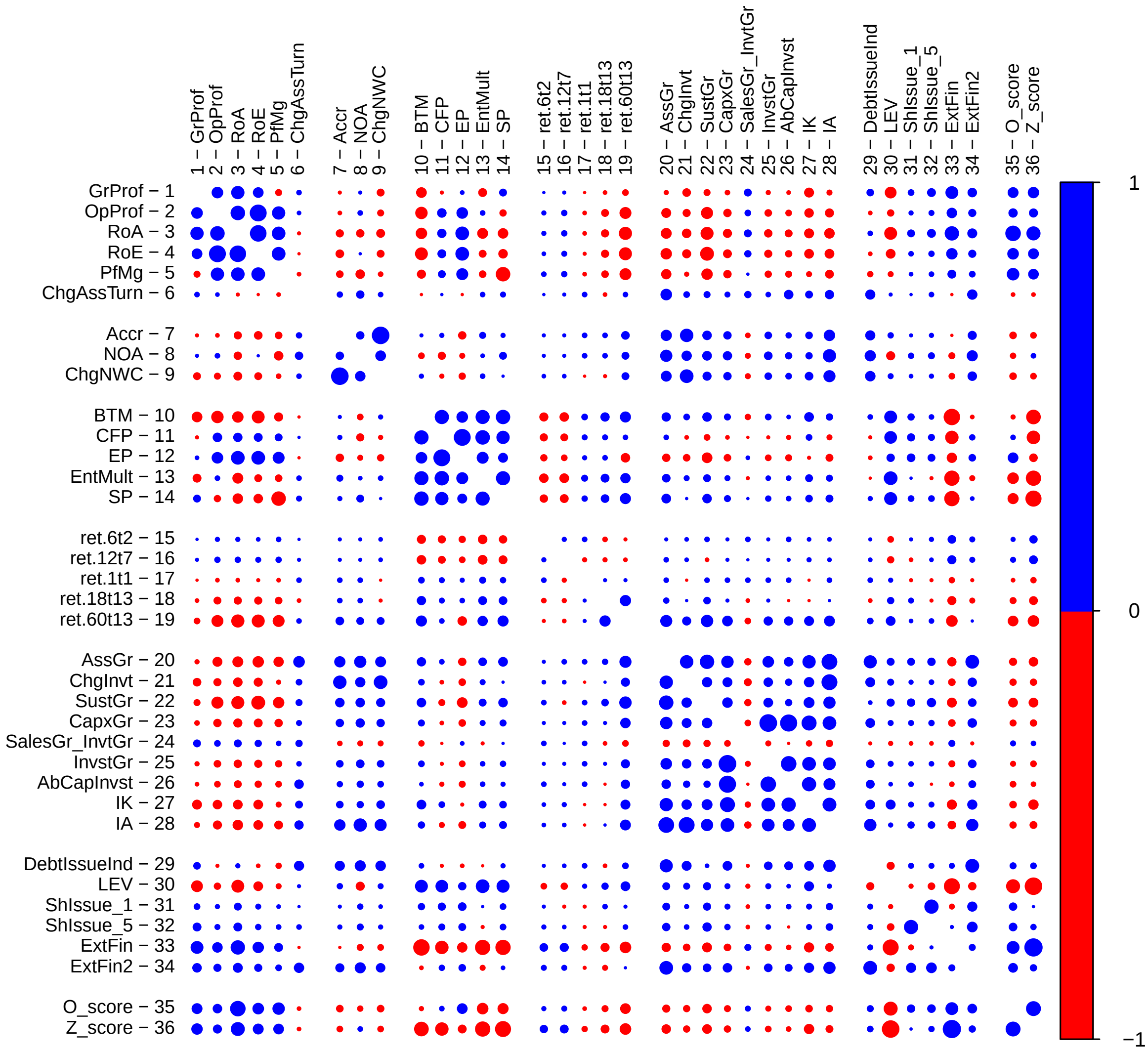

Figure 2: A matrix heat map of pairwise correlations among the 36 anomaly signals. Blue circles represent positive correlation while red circles are negative correlations. A bigger circle represents higher magnitude of correlations. Lower half is symmetric to the upper half. All correlations are calculated for the entire sample.

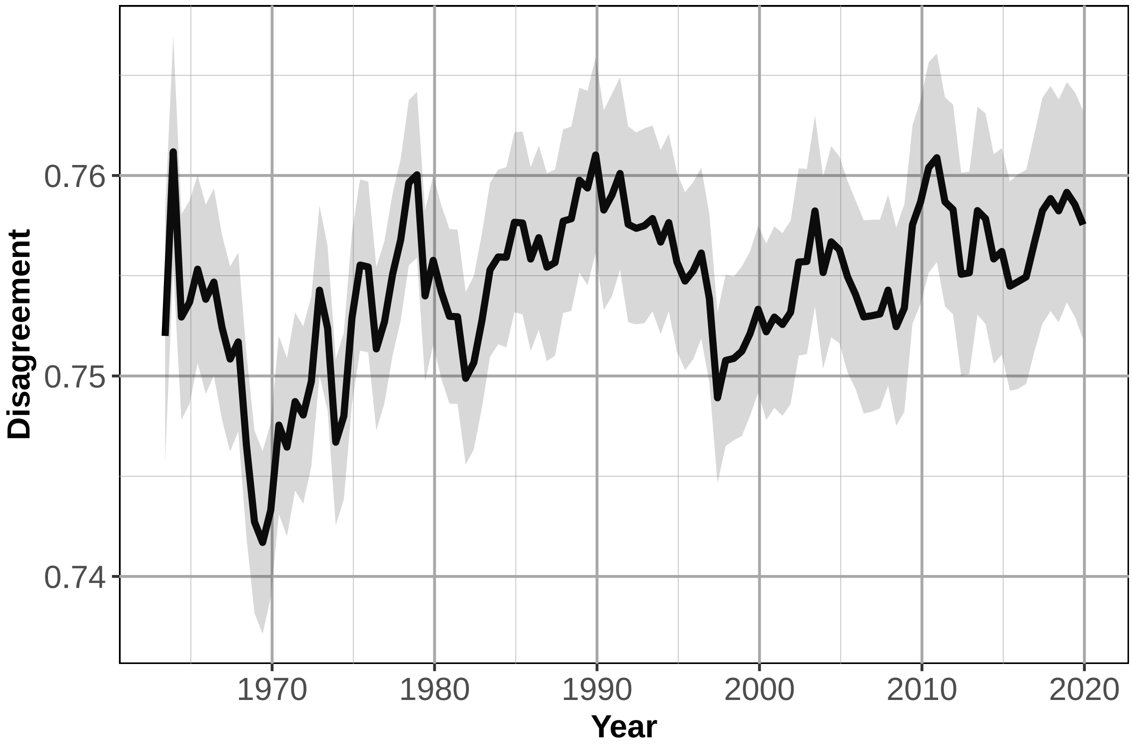

Figure 3: Monthly cross-sectional mean and standard error of the disagreement measure from year 1966. The confidence interval is set to two standard errors.

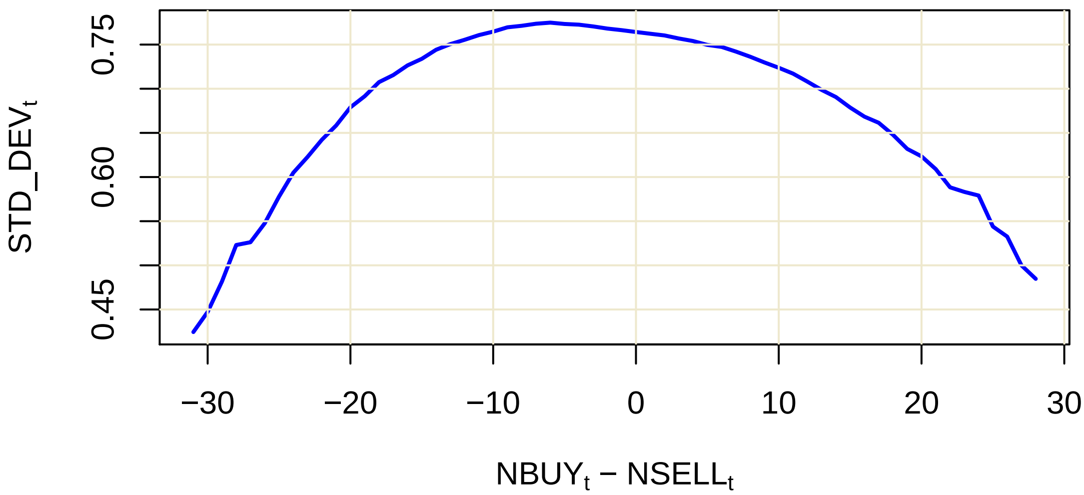

Figure 4: Disagreement averaged over entore NYSE stocks (1962-2019) against the difference between number of buy and sell signals.

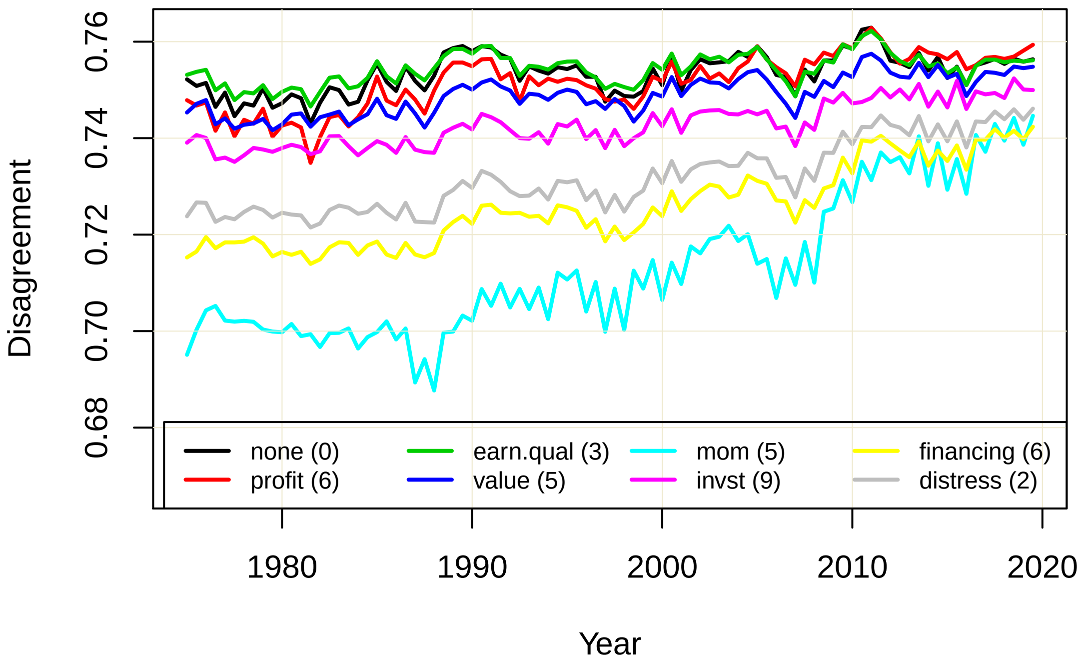

Figure 5: Average (semi-annual) Disagreement, since 1975, using all but one group of anomalies (see Table 1) over time. Disagreement is the standard deviation of trading signals generated from anomalies in equation 2.

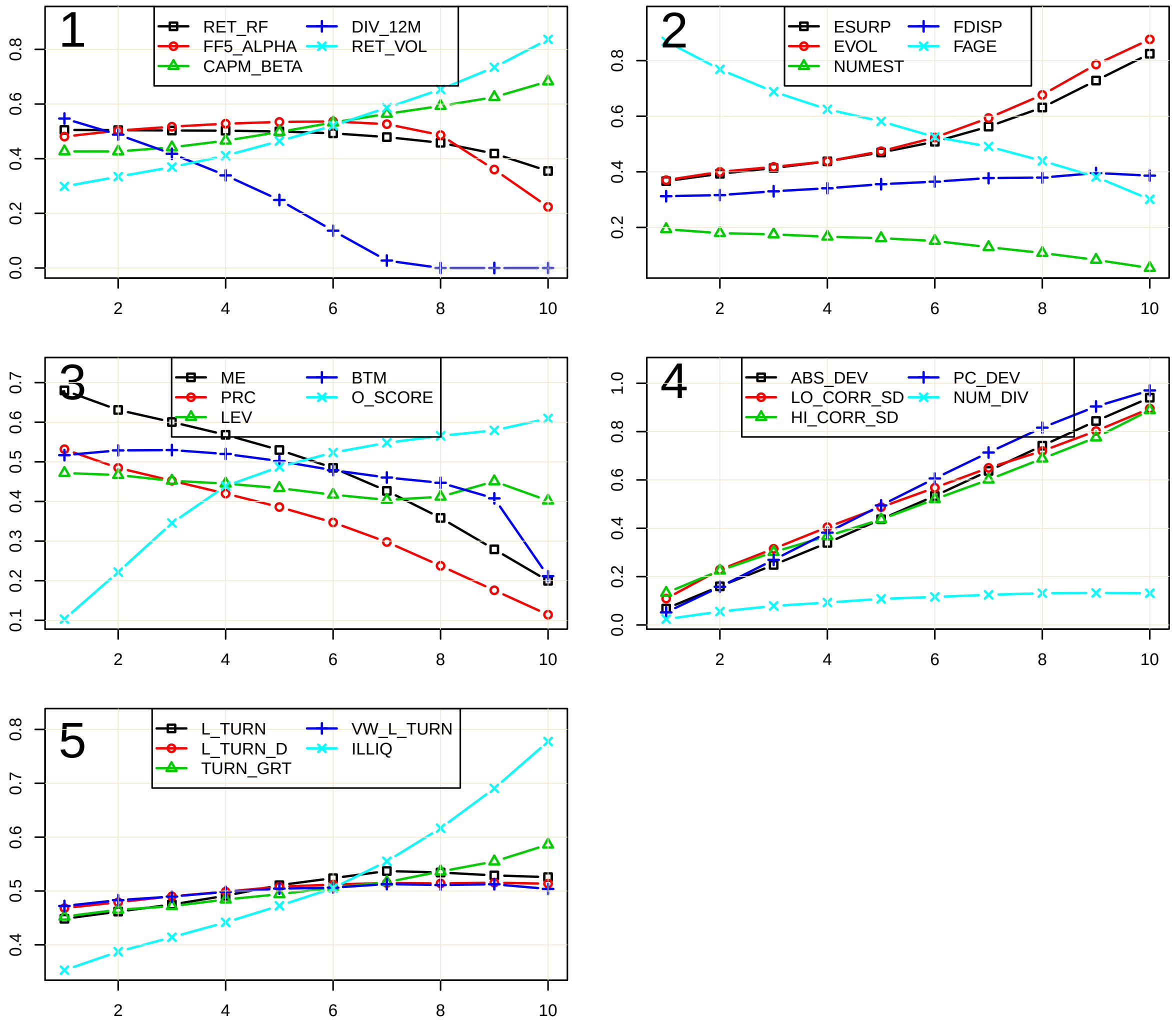

Figure 6: Cross-sectional ranks of firm characteristics across disagreement deciles. Ranks appear on vertical axis and disagreement decile on horizontal axis. Both ranks and disagreement deciles are constructed each month. Out of 5 subplots, first 3 have ranks of firm characteristics, next is the ranks of different disagreement measure and the last subplot has different turnover measures. Construction details of all variables appear in Appendix A.2

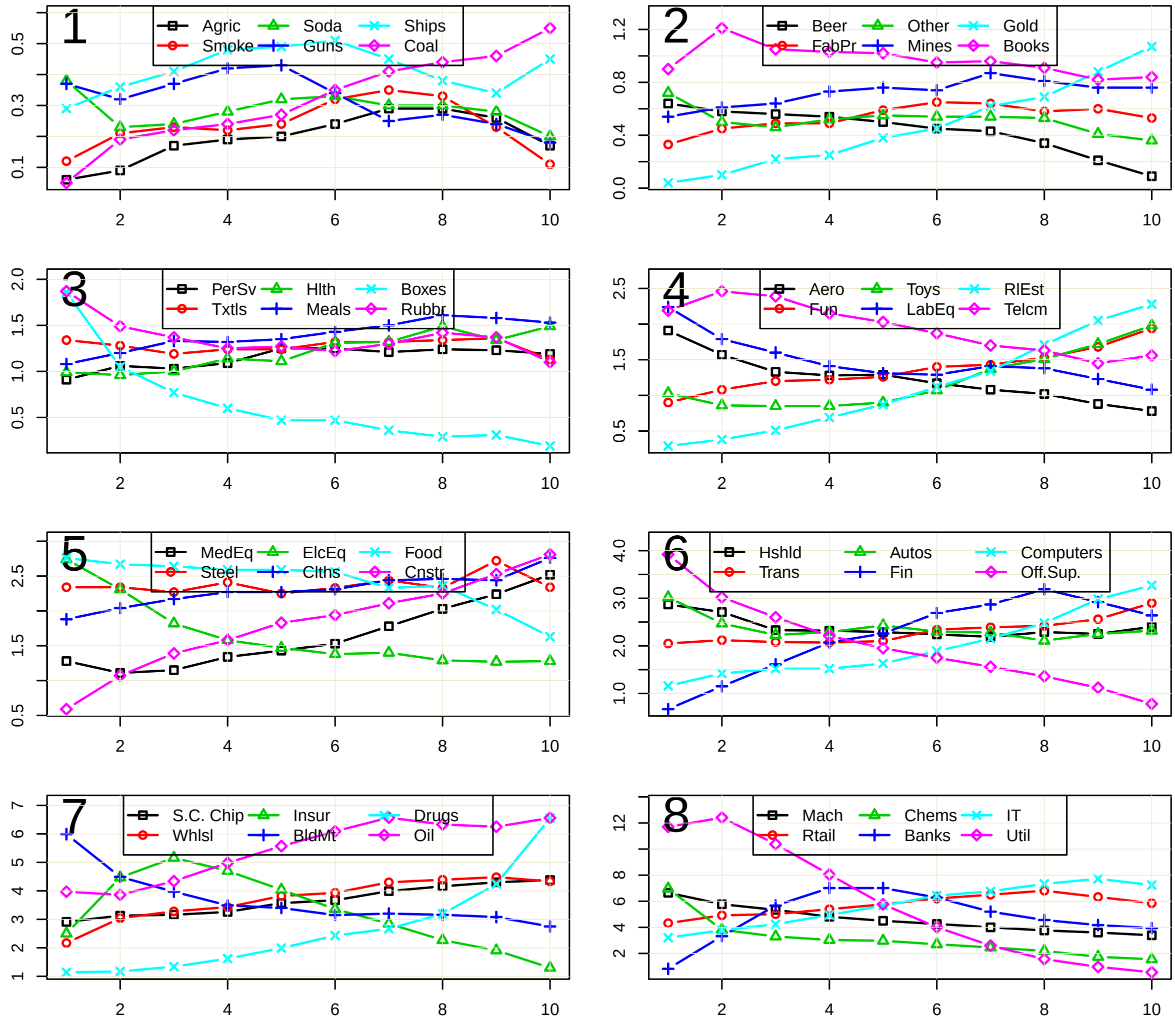

Figure 7: Relative industry concentration across disagreement deciles. For each disagreement decile (horizontal axis) the percentage of firm-month observations belonging to a particular industry appear on vertical axis. Both percentage of firm-month observations and disagreement deciles are calculated cross-sectionally for each month. 48 Fama and French (1997) industries are presented in 8 subplots of 6 industries each. Industries are sorted on their relative concentration (vertical axis) and then the smallest group is presented in subplot 1 and the largest in subplot 8.

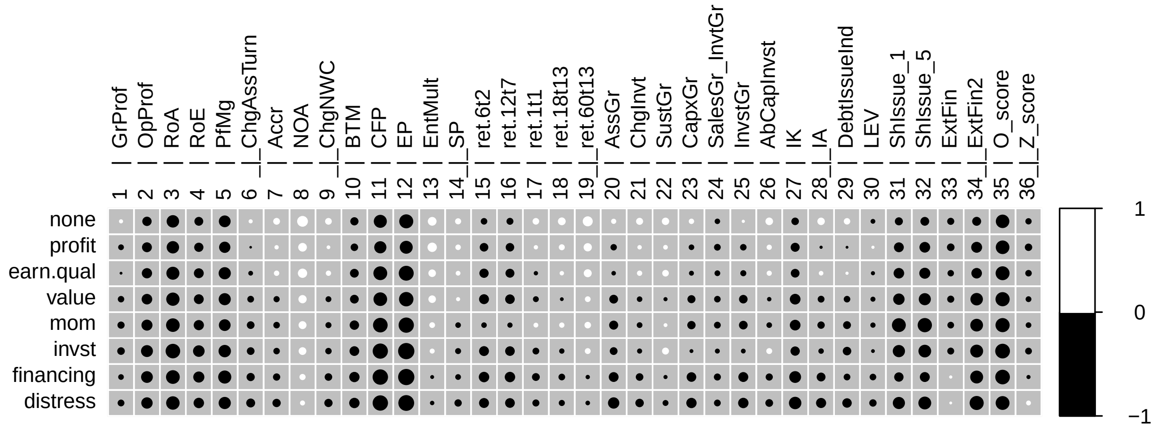

Figure 8: Correlation heat map of 36 signals and disagreement. The row names gives the category of signals left out in the disagreement measure (see Table 1). Positive correlation is represented by white circles while negative correlation is shown by black circles. The size of the circle is proportional to the magnitude of the correlation coefficient.

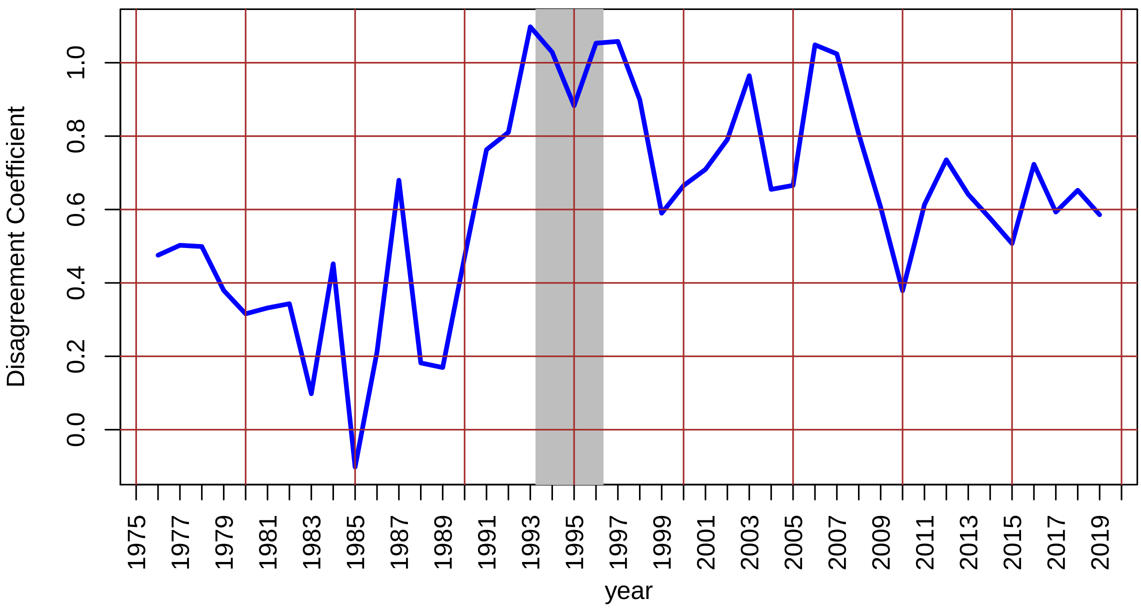

Figure 9: Disagreement coefficient from annual regressions. The EDGAR implementation period (April 1993 to May 1996) is shown in shaded region.

Table 1: **Anomalies List**

Category / S.No.	Anomaly	Predicted Relationship	Source
Profitability			
1	Gross Profitability	Positive	Novy-Marx (2013)
2	Operating Profitability	Positive	Fama and French (2015)
3	Return on Assets	Positive	Haugen et al. (1996)
4	Return on Equity	Positive	Haugen et al. (1996)
5	Profit Margin	Positive	Soliman (2008)
6	Change in Asset Turnover	Positive	Soliman (2008)
Earnings Quality			
7	Accruals	Negative	Sloan (1996)
8	Net Operating Assets	Negative	Hirshleifer et al. (2004)
9	Changes in Net Working Capital	Negative	Soliman (2008)
Valuation			
10	Book to market	Positive	Fama and French (1992)
11	Cash flow to price	Positive	Lakonishok et al. (1994)
12	Earnings to Price	Positive	Basu (1977)
13	Enterprise Multiple	Negative	Loughran and Wellman (2011)
14	Sales to price	Positive	Barbee Jr et al. (1996)
Momentum			
15	Short term momentum	Positive	Jegadeesh and Titman (1993)
16	Lagged Momentum	Positive	Novy-Marx (2012)
17	Short-term reversal	Negative	Jegadeesh and Titman (1993)
18	Medium-term reversal	Negative	Jegadeesh and Titman (1993)
19	Long-term reversal	Negative	De Bondt and Thaler (1985)
Investment			
20	Asset Growth	Negative	Cooper et al. (2008)
21	Inventory Growth	Negative	Thomas and Zhang (2002)
22	Sustainable Growth	Negative	Lockwood and Prombutr (2010)
23	CAPX Growth	Negative	Abarbanell and Bushee (1998)
24	Growth in Sales minus growth in Inventory	Positive	Abarbanell and Bushee (1998)
25	Investment growth	Negative	Xing (2007)
26	Abnormal CAPX	Negative	Titman et al. (2004)
27	Investment to Capital Ratio	Negative	Xing (2007)
28	Investment to Asset Ratio	Negative	Lyandres et al. (2007)
Financing			
29	Increase in Debt Issuance	Negative	Spiess and Affleck-Graves (1999)
30	Leverage	Positive	Bhandari (1988)
31	One year Share Issuance	Negative	Pontiff and Woodgate (2008)
32	Five year Share Issuance	Negative	Daniel and Titman (2006)
33	External Financing – I	Negative	Bradshaw et al. (2006)
34	External Financing – II	Negative	Bradshaw et al. (2006)
Distress			
35	O-Score	Negative	Dichev (1998)
36	Z-Score	Positive	Dichev (1998)

List of 36 anomalies used to construct the disagreement measure. 31 anomalies are from Linnainmaa and Roberts (2018) and 5 momentum anomalies from McLean and Pontiff (2016). Predicted relationship is from the original study findings.

Table 2: **Correlations among explanatory variables**

	Time-series average of cross-sectional correlation coefficients												
	RET^+	RET^-	LEV	$CAPM\_\beta$	BTM	L_PRC	L_FAGE	L_ME	$ESURP$	$EVOL$	$NUMEST$	$FDISP$	STD_DEV
RET^+		0.729	-0.043	0.043	-0.089	0.090	-0.005#	0.053	0.022	0.005#	-0.003#	0.004#	0.009
RET^-	0.273		-0.039	-0.092	-0.076	0.182	0.066	0.133	-0.066	-0.08	0.052	-0.042	-0.123
LEV	-0.004	-0.086		0.017	0.406	-0.172	0.096	-0.057	0.254	0.369	-0.062	0.14	-0.006
$CAPM\_\beta$	0.087	-0.130	0.025		-0.071	-0.153	-0.177	-0.074	0.144	0.114	-0.022	0.169	0.174
BTM	-0.025	-0.053	0.304	-0.013		-0.263	0.119	-0.265	0.277	0.396	-0.197	0.108	-0.089
L_PRC	-0.047	0.270	-0.177	-0.14	-0.076		0.325	0.789	-0.444	-0.575	0.538	-0.051	-0.379
L_FAGE	-0.041	0.082	0.008	-0.145	0.033	0.193		0.361	-0.032	-0.032	0.201	-0.017	-0.257
L_ME	-0.066	0.201	-0.137	-0.065	-0.12	0.779	0.237		-0.156	-0.165	0.772	-0.037	-0.338
$ESURP$	0.046	-0.071	0.227	0.06	0.033	-0.217	-0.096	-0.389		0.661	-0.237	0.185	0.279
$EVOL$	0.042	-0.071	0.226	0.063	-0.009#	-0.231	-0.068	-0.495	0.637		-0.318	0.188	0.310
$NUMEST$	-0.065	0.078	-0.079	-0.047	-0.105	0.468	0.228	0.810	-0.089	-0.102		0.318	-0.148
$FDISP$	0.014	-0.043	0.049	0.067	0.06	-0.098	-0.014	0.133	0.070	0.06	-0.001#		0.042
STD_DEV	0.099	-0.184	0.135	0.18	0.026	-0.393	-0.215	-0.337	0.149	0.163	-0.118	0.077	

At each month, cross-sectional correlation among variables is computed and then their time-series average over the duration of sample is reported. Lower triangle represents Pearson correlation while the upper triangle consists of Spearman rank correlations.

[*] Correlations having a trailing '#' are NOT significant at the 5% level.

Table 3: **Descriptive Statistics**

	Descriptive Statistics of explanatory variables, measures of turnover and measures of disagreement										
	Mean	**SD**	**Min**	**p25**	**Median**	**p75**	**Max**	**IQR***	**Range***	**Skew**	**Kurt**
RET_Rf	0.008	0.056	-0.286	-0.022	0.009	0.039	0.311	1.099	10.723	-0.157	6.429
RET^+	0.050	0.033	0.001	0.027	0.044	0.064	0.322	1.115	9.685	2.074	12.491
RET^-	-0.039	0.029	-0.281	-0.048	-0.032	-0.021	-0.002	0.953	9.509	-2.715	15.712
LEV	0.001	0.001	0.000	0.001	0.001	0.001	0.006	0.861	9.811	3.279	21.491
$CAPM\_\beta$	1.103	0.190	0.631	1.010	1.146	1.249	1.388	1.257	3.975	-0.683	2.632
BTM	0.001	0.000	-0.001	0.001	0.001	0.001	0.003	1.081	9.687	1.353	7.431
L_PRC	2.754	0.288	1.771	2.576	2.760	2.910	3.631	1.164	6.465	-0.070	3.409
L_FAGE	4.513	0.648	0.000	4.485	4.717	4.863	5.067	0.584	7.816	-3.295	16.970
L_ME	12.249	1.309	9.961	10.987	12.098	13.409	14.542	1.850	3.499	0.096	1.755
$ESURP$	0.151	0.715	0.007	0.029	0.042	0.078	11.194	0.068	15.650	12.258	166.757
$EVOL$	0.164	0.448	0.006	0.034	0.055	0.115	3.996	0.180	8.912	5.904	39.443
$NUMEST$	8.674	1.365	5.063	8.015	8.557	9.646	11.302	1.195	4.571	-0.405	2.832
$FDISP$	0.156	0.071	0.038	0.099	0.147	0.201	0.402	1.444	5.150	0.622	2.997
DIV_12M	0.021	0.008	0.007	0.014	0.018	0.026	0.041	1.484	4.185	0.663	2.356
STD_DEV	0.755	0.004	0.739	0.753	0.755	0.758	0.763	1.217	6.138	-0.996	4.347
STD_DEV_R	0.474	0.014	0.440	0.464	0.473	0.484	0.507	1.457	4.891	0.038	2.332
ABS_DEV	0.625	0.004	0.608	0.624	0.626	0.628	0.631	1.003	6.068	-1.444	5.510
PC_DEV	0.761	0.005	0.746	0.758	0.760	0.764	0.771	1.282	5.125	-0.144	2.591
NUM_FLIPS	0.125	0.418	0.000	0.006	0.008	0.012	2.501	0.015	5.982	3.551	14.480
NUM_DIV	4.941	3.841	0.000	3.902	4.197	4.479	21.403	0.150	5.572	2.906	10.762

Table 3: **Descriptive Statistics** *(continued)*

	Mean	**SD**	**Min**	**p25**	**Median**	**p75**	**Max**	**IQR***	**Range***	**Skew**	**Kurt**
L_TURN	-0.886	0.786	-2.461	-1.457	-0.980	-0.170	0.648	1.638	3.958	0.129	2.004
L_TURN_R	0.500	0.000	0.498	0.500	0.500	0.500	0.500	0.635	9.735	-2.978	15.986
L_TURN_1d	0.494	0.868	-1.076	-0.186	0.331	1.318	2.113	1.733	3.675	0.272	1.892
L_TURN_5d	2.196	0.822	0.569	1.618	2.096	2.957	3.860	1.629	4.003	0.143	1.996
L_TURN_D	0.004	0.237	-0.768	-0.145	-0.019	0.134	0.818	1.176	6.684	0.238	3.559
L_TURN_GRT	-2.509	0.372	-4.019	-2.720	-2.500	-2.322	-1.262	1.070	7.421	0.078	3.943
L_TURN_ILLIQ	0.002	0.177	-0.327	-0.119	-0.030	0.089	0.978	1.174	7.366	1.338	6.259
VW_L_TURN	0.005	0.150	-0.442	-0.103	0.004	0.105	0.443	1.388	5.918	0.026	3.050
EW_L_TURN	-0.002	0.080	-0.542	-0.037	0.007	0.044	0.197	1.010	9.236	-1.615	9.754

At each month the average of all the variables is computed and their time-series descriptive statistics are reported. Variable definitions are present in Appendix A.2

* Interquartile range (IQR) and range (Max - Min) are in multiples of standard deviation.

Table 4: **Monthly cross-sectional regression: different specifications**

	L_TURN_t					
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.407***	1.440***	1.319***	1.325***	1.346***	1.231***
	(0.098)	(0.099)	(0.092)	(0.094)	(0.094)	(0.088)
RET^{-}_{t-1}	-1.814***	-1.816***	-1.926***	-1.709***	-1.708***	-1.813***
	(0.117)	(0.118)	(0.120)	(0.112)	(0.113)	(0.114)
LEV_{t-1}	36.754***	42.017***	39.707***	31.733***	35.144***	34.274***
	(8.239)	(8.065)	(9.037)	(7.944)	(7.698)	(8.709)
$CAPM\_\beta_{t-1}$	0.410***	0.412***	0.451***	0.396***	0.397***	0.436***
	(0.017)	(0.017)	(0.019)	(0.017)	(0.017)	(0.019)
BTM_{t-1}	9.880	15.147	9.429	17.883	21.898*	17.828
	(12.847)	(12.596)	(13.284)	(12.756)	(12.519)	(13.166)
L_PRC_{t-1}	0.188***	0.183***	0.293***	0.205***	0.202***	0.311***
	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)
L_FAGE_{t-1}	-0.033**	-0.033**	-0.002	-0.023	-0.023	0.009
	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.018)
$ESURP_{t-1}$	0.181***		0.248***	0.151***		0.214***
	(0.054)		(0.058)	(0.050)		(0.054)
$EVOL_{t-1}$	0.079		0.093	0.032		0.043
	(0.063)		(0.069)	(0.059)		(0.063)
$NUMEST_{t-1}$	0.026***	0.026***		0.026***	0.026***	
	(0.002)	(0.002)		(0.002)	(0.002)	
$FDISP_{t-1}$	0.109***	0.111***		0.099***	0.101***	
	(0.011)	(0.011)		(0.010)	(0.010)	
STD_DEV_{t-1}				0.736***	0.749***	0.785***
				(0.070)	(0.070)	(0.073)
Within R^2	0.198	0.198	0.149	0.205	0.204	0.157
Adj. R^2	0.535	0.535	0.507	0.539	0.539	0.511
$\%R^2$ Explained				0.78	0.80	0.99
Observations	399,476	399,476	399,476	399,476	399,476	399,476

Log turnover regressed on different set of explanatory variables. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 5: **Monthly cross-sectional regression: daily, weekly and monthly turnover**

	$L\_TURN\_1d_t$	$L\_TURN\_5d_t$	L_TURN_t	$L\_TURN\_1d_t$	$L\_TURN\_5d_t$	L_TURN_t
	(1)	(2)	(3)	(4)	(5)	(6)
RET^+_{t-1}	2.416***	1.987***	1.380***	2.344***	1.907***	1.297***
	(0.147)	(0.119)	(0.096)	(0.144)	(0.115)	(0.092)
RET^-_{t-1}	-2.854***	-2.369***	-1.787***	-2.761***	-2.267***	-1.680***
	(0.144)	(0.117)	(0.116)	(0.142)	(0.113)	(0.111)
LEV_{t-1}	38.074***	37.480***	38.087***	33.665***	32.604***	33.006***
	(8.371)	(8.210)	(8.265)	(8.121)	(7.915)	(7.937)
$CAPM\_\beta_{t-1}$	0.416***	0.408***	0.399***	0.404***	0.395***	0.386***
	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)
BTM_{t-1}	-1.200	5.515	10.433	5.940	13.412	18.662
	(13.080)	(13.031)	(12.832)	(13.011)	(12.937)	(12.737)
L_PRC_{t-1}	0.247***	0.213***	0.183***	0.262***	0.229***	0.200***
	(0.017)	(0.017)	(0.016)	(0.018)	(0.017)	(0.016)
L_FAGE_{t-1}	-0.015	-0.029*	-0.034**	-0.006	-0.019	-0.024
	(0.017)	(0.017)	(0.016)	(0.018)	(0.017)	(0.016)
$ESURP_{t-1}$	0.155**	0.212***	0.183***	0.129**	0.183***	0.153***
	(0.062)	(0.057)	(0.055)	(0.059)	(0.053)	(0.051)
$EVOL_{t-1}$	-0.018	0.002	0.070	-0.059	-0.044	0.022
	(0.063)	(0.068)	(0.063)	(0.061)	(0.065)	(0.059)
$NUMEST_{t-1}$	0.029***	0.027***	0.026***	0.029***	0.027***	0.025***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$FDISP_{t-1}$	0.101***	0.104***	0.105***	0.093***	0.095***	0.095***
	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
STD_DEV_{t-1}				0.644***	0.712***	0.742***
				(0.073)	(0.071)	(0.070)
Within R^2	0.172	0.193	0.193	0.175	0.198	0.199
Adj. R^2	0.494	0.517	0.534	0.496	0.520	0.538
$\% R^2$ Explained				0.38	0.64	0.81
Observations	396,565	396,565	396,565	396,565	396,565	396,565

Log turnover, measured over daily, weekly and monthly intervals, is regressed on a set of explanatory variables and the disagreement measure STD_DEV. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 6: **Monthly cross-sectional regression: different measures of turnover**

	L_TURN_t	ΔL_TURN_t	$L\_TURN\_GRT_t$	$L\_TURN\_D_t$	$L\_TURN\_ILLIQ_t$	$VW\_L\_TURN_t$	$EW\_L\_TURN_t$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RET^{+}_{t-1}	1.324***	-0.888***	1.136***	1.064***	0.682***	1.019***	0.974***
	(0.094)	(0.056)	(0.101)	(0.064)	(0.037)	(0.068)	(0.061)
RET^{-}_{t-1}	-1.709***	0.747***	-0.995***	-1.247***	-1.409***	-1.187***	-1.365***
	(0.112)	(0.089)	(0.171)	(0.094)	(0.070)	(0.072)	(0.057)
LEV_{t-1}	31.780***	0.147	17.815**	0.504	14.304***	-0.296	-0.964
	(7.989)	(0.868)	(6.997)	(2.844)	(4.425)	(2.801)	(2.893)
$CAPM\_\beta_{t-1}$	0.396***	0.010***	-0.037**	0.018**	0.007	0.037***	0.024***
	(0.017)	(0.002)	(0.018)	(0.007)	(0.007)	(0.007)	(0.007)
BTM_{t-1}	17.781	4.046*	2.406	-12.665**	39.128***	-12.828**	-9.964*
	(12.781)	(2.219)	(12.200)	(5.295)	(7.474)	(5.124)	(5.339)
L_PRC_{t-1}	0.205***	-0.012***	-0.124***	0.075***	-0.050***	0.063***	0.073***
	(0.017)	(0.003)	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)
L_FAGE_{t-1}	-0.023	-0.003***	0.171***	-0.017***	0.020***	0.021***	-0.000
	(0.017)	(0.001)	(0.021)	(0.006)	(0.007)	(0.006)	(0.006)
$ESURP_{t-1}$	0.154***	0.028	0.022	0.124***	0.105**	0.063*	0.092***
	(0.050)	(0.019)	(0.065)	(0.032)	(0.041)	(0.032)	(0.030)
$EVOL_{t-1}$	0.032	0.051**	0.066	-0.052	0.140***	0.003	-0.003
	(0.059)	(0.025)	(0.087)	(0.050)	(0.050)	(0.044)	(0.045)
$NUMEST_{t-1}$	0.026***	0.000**	-0.033***	-0.000	0.005***	-0.000	0.003***
	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$FDISP_{t-1}$	0.099***	-0.002	0.007	0.005	0.038***	0.007	0.011**
	(0.010)	(0.002)	(0.013)	(0.005)	(0.006)	(0.005)	(0.005)
STD_DEV_{t-1}	0.737***	-0.006	0.517***	0.241***	0.090***	0.245***	0.183***
	(0.070)	(0.008)	(0.080)	(0.032)	(0.030)	(0.033)	(0.033)
Within R^2	0.205	0.018	0.024	0.034	0.077	0.031	0.037
Adj. R^2	0.539	0.020	0.048	0.136	0.184	0.062	0.053
$\%R^2$ Explained	0.78	-0.00	0.08	0.17	-0.03	0.16	0.06
Observations	399,409	399,409	399,409	399,409	399,409	399,409	399,409

Several measures of tunover are regressed on a set of explanatory variables and the disagreement measure STD_DEV. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. $\%$ R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 7: **Monthly cross-sectional regression: different measures of disagreement**

	L_TURN_t				
	(1)	(2)	(3)	(4)	(5)
STD_DEV_{t-1}	0.674***				
	(0.078)				
ABS_DEV_{t-1}		0.595***			
		(0.065)			
PC_DEV_{t-1}			0.747***		
			(0.081)		
$LO\_CORR\_SD_{t-1}$				0.232***	
				(0.057)	
$HI\_CORR\_SD_{t-1}$				0.377***	
				(0.041)	
NUM_FLIPS_{t-1}					0.003*
					(0.002)
NUM_DIV_{t-1}					0.001***
					(0.000)
Within R^2	0.194	0.195	0.195	0.194	0.188
Adj. R^2	0.548	0.548	0.548	0.547	0.544
$\% R^2$ Explained	0.74	0.82	0.79	0.76	0.01
Observations	297,725	297,725	297,725	297,725	297,725

Log turnover regressed on set of controls and several measures of disagreement. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 8: **Monthly cross-sectional regression: rank regression**

	L_TURN_t		$L\_TURN\_R_t$	
	(1)	(2)	(3)	(4)
$RET^{+}\_R_{t-1}$	0.332***	0.310***	0.101***	0.093***
	(0.011)	(0.011)	(0.003)	(0.003)
$RET^{-}\_R_{t-1}$	-0.368***	-0.339***	-0.124***	-0.113***
	(0.013)	(0.012)	(0.003)	(0.003)
LEV_R_{t-1}	0.368***	0.374***	0.109***	0.112***
	(0.038)	(0.038)	(0.011)	(0.011)
$CAPM\_\beta\_R_{t-1}$	0.702***	0.686***	0.215***	0.210***
	(0.031)	(0.031)	(0.009)	(0.009)
BTM_R_{t-1}	-0.248***	-0.194***	-0.074***	-0.056***
	(0.041)	(0.042)	(0.012)	(0.012)
$L\_PRC\_R_{t-1}$	0.687***	0.729***	0.198***	0.212***
	(0.055)	(0.055)	(0.016)	(0.016)
$L\_FAGE\_R_{t-1}$	-0.056	-0.040	-0.021*	-0.015
	(0.035)	(0.035)	(0.011)	(0.011)
$ESURP\_R_{t-1}$	0.136***	0.118***	0.048***	0.042***
	(0.015)	(0.015)	(0.005)	(0.005)
$EVOL\_R_{t-1}$	0.184***	0.138***	0.069***	0.054***
	(0.034)	(0.034)	(0.010)	(0.011)
$NUMEST\_R_{t-1}$	0.840***	0.837***	0.275***	0.274***
	(0.057)	(0.057)	(0.018)	(0.018)
$FDISP\_R_{t-1}$	0.475***	0.476***	0.141***	0.141***
	(0.026)	(0.026)	(0.007)	(0.007)
$STD\_DEV\_R_{t-1}$		0.239***		0.081***
		(0.027)		(0.008)
Within R^2	0.237	0.241	0.236	0.241
Adj. R^2	0.558	0.560	0.292	0.296
$\%R^2$ Explained		0.52		1.24
Observations	399,476	399,476	399,476	399,476

Cross-sectional rank of log turnover regressed on cross-sectional ranks of set of controls and several measures of disagreement. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 9: **Monthly cross-sectional regression: size terciles**

	L_TURN_t					
	SMALL		*MEDIUM*		*BIG*	
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.471***	1.378***	1.469***	1.340***	1.443***	1.370***
	(0.134)	(0.127)	(0.111)	(0.105)	(0.086)	(0.084)
RET^{-}_{t-1}	-1.743***	-1.631***	-1.603***	-1.441***	-1.913***	-1.805***
	(0.113)	(0.108)	(0.134)	(0.124)	(0.136)	(0.132)
LEV_{t-1}	16.856**	12.782*	37.760**	27.482**	114.072***	105.265***
	(7.405)	(7.047)	(14.776)	(12.989)	(18.340)	(17.713)
$CAPM\_\beta_{t-1}$	0.313***	0.295***	0.337***	0.313***	0.298***	0.284***
	(0.028)	(0.028)	(0.021)	(0.021)	(0.024)	(0.024)
BTM_{t-1}	21.952**	25.090**	17.529	32.060	67.983**	95.846***
	(10.668)	(10.362)	(22.050)	(21.255)	(29.171)	(28.935)
L_PRC_{t-1}	0.206***	0.251***	0.103***	0.130***	0.018	0.018
	(0.024)	(0.023)	(0.025)	(0.025)	(0.022)	(0.022)
L_FAGE_{t-1}	-0.049	-0.041	0.026	0.036*	-0.066***	-0.054**
	(0.032)	(0.032)	(0.020)	(0.019)	(0.022)	(0.022)
$ESURP_{t-1}$	0.151***	0.130***	0.259***	0.208***	0.465***	0.424***
	(0.052)	(0.047)	(0.074)	(0.063)	(0.095)	(0.091)
$EVOL_{t-1}$	0.311***	0.251***	0.197*	0.133	0.013	-0.110
	(0.070)	(0.063)	(0.120)	(0.096)	(0.368)	(0.325)
$NUMEST_{t-1}$	0.140***	0.140***	0.059***	0.059***	0.013***	0.012***
	(0.009)	(0.009)	(0.003)	(0.003)	(0.002)	(0.002)
$FDISP_{t-1}$	0.041***	0.031**	0.051***	0.035***	0.040**	0.028*
	(0.015)	(0.015)	(0.011)	(0.011)	(0.016)	(0.016)
STD_DEV_{t-1}		1.351***		1.129***		0.579***
		(0.148)		(0.088)		(0.083)
Within R^2	0.203	0.215	0.195	0.212	0.134	0.140
Adj. R^2	0.429	0.438	0.600	0.608	0.633	0.636
$\%R^2$ Explained		0.88		1.16		0.12
Observations	56,292	56,292	175,401	175,401	167,783	167,783

Log turnover regressed on set of controls and several measures of disagreement across three size terciles based on 70/30 NYSE breakpoints. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 10: **Monthly cross-sectional regression: 5-year subperiods**

	L_TURN_t								
	1975-1979	*1980-1984*	*1985-1989*	*1990-1994*	*1995-1999*	*2000-2004*	*2005-2009*	*2010-2014*	*2015-2019*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RET^{+}_{t-1}	2.359***	1.719***	1.784***	1.735***	1.284***	1.148***	0.751***	1.066***	0.843***
	(0.177)	(0.186)	(0.167)	(0.156)	(0.125)	(0.104)	(0.157)	(0.110)	(0.129)
RET^{-}_{t-1}	-0.579*	-0.698**	-0.811**	-1.877***	-2.056***	-2.169***	-1.919***	-2.020***	-1.776***
	(0.325)	(0.295)	(0.355)	(0.223)	(0.187)	(0.138)	(0.271)	(0.175)	(0.136)
LEV_{t-1}	38.657***	19.221***	36.408***	-0.118	21.033	3.770	2.020***	7.018**	2.680*
	(10.802)	(3.191)	(10.549)	(2.469)	(15.956)	(8.272)	(0.588)	(2.790)	(1.527)
$CAPM\_\beta_{t-1}$	0.530***	0.492***	0.524***	0.610***	0.423***	0.503***	0.382***	0.386***	0.273***
	(0.059)	(0.046)	(0.051)	(0.056)	(0.038)	(0.041)	(0.030)	(0.028)	(0.032)
BTM_{t-1}	154.316***	32.619	43.169**	3.571	76.326***	30.185*	14.198**	35.722**	36.779**
	(34.574)	(24.722)	(18.011)	(3.427)	(29.547)	(15.448)	(7.155)	(16.594)	(14.653)
L_PRC_{t-1}	0.039	0.074*	0.076**	0.152***	0.183***	0.403***	0.342***	0.171***	0.099***
	(0.043)	(0.040)	(0.038)	(0.035)	(0.030)	(0.039)	(0.034)	(0.022)	(0.021)
L_FAGE_{t-1}	0.105	0.173***	0.201***	0.007	-0.020	-0.022	-0.066**	-0.097***	-0.113***
	(0.064)	(0.048)	(0.041)	(0.038)	(0.030)	(0.026)	(0.026)	(0.026)	(0.024)
$ESURP_{t-1}$	-0.005	0.120	-0.047*	0.011	0.006	-0.024	0.001	0.031	-0.018
	(0.079)	(0.087)	(0.024)	(0.015)	(0.053)	(0.015)	(0.012)	(0.047)	(0.016)
$EVOL_{t-1}$	-0.198	0.106	-0.017	0.012	-0.003	0.086	0.129**	0.137	0.057***
	(0.140)	(0.147)	(0.037)	(0.016)	(0.148)	(0.135)	(0.059)	(0.109)	(0.012)
$NUMEST_{t-1}$	0.004	0.028***	0.027***	0.025***	0.027***	0.028***	0.030***	0.035***	0.024***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)
$FDISP_{t-1}$	0.040*	0.008**	0.010**	0.011	0.029***	0.020***	0.047***	0.027***	0.016***
	(0.023)	(0.004)	(0.005)	(0.007)	(0.008)	(0.007)	(0.015)	(0.006)	(0.005)
STD_DEV_{t-1}	0.511***	0.401***	0.317**	0.930***	1.038***	0.907***	1.061***	0.761***	0.694***
	(0.149)	(0.131)	(0.142)	(0.155)	(0.145)	(0.144)	(0.138)	(0.139)	(0.123)
Within R^2	0.130	0.130	0.205	0.209	0.213	0.295	0.257	0.269	0.167
Adj. R^2	0.227	0.243	0.282	0.283	0.301	0.387	0.386	0.351	0.280
% R^2 Explained	0.24	0.48	0.40	0.40	0.60	1.48	1.71	0.76	0.38
Observations	24,030	41,553	38,056	39,468	47,455	50,244	53,200	53,261	52,209

Log turnover regressed on set of controls and several measures of disagreement over nine 5-year subperiods. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 11: **Univariate Sorts: Different explanatory variables**

	L_TURN_t									
	D_1	D_2-D_1	D_3-D_2	D_4-D_3	D_5-D_4	D_6-D_5	D_7-D_6	D_8-D_7	D_9-D_8	$D_{10}-D_9$
RET^{+}_{t-1}	-1.051	-0.002	0.027**	0.031**	0.048**	0.060**	0.045**	0.080**	0.086**	0.166**
RET^{-}_{t-1}	-0.767	-0.137**	-0.046**	-0.054**	-0.038**	-0.030**	-0.026**	-0.016*	-0.006	0.188**
LEV_{t-1}	-0.985	0.133**	0.002	0.036**	0.000	-0.004	-0.029**	-0.018**	-0.030**	0.140**
$CAPM\_\beta_{t-1}$	-1.410	0.169**	0.132**	0.116**	0.102**	0.089**	0.093**	0.097**	0.098**	0.148**
BTM_{t-1}	-0.686	0.036**	-0.081**	-0.078**	-0.069**	-0.054**	-0.068**	-0.023**	-0.010	0.185**
L_PRC_{t-1}	-1.485	0.238**	0.186**	0.065**	0.077**	0.094**	0.143**	0.160**	0.125**	0.019**
L_FAGE_{t-1}	-0.714	-0.027**	0.005	-0.037**	-0.069**	-0.155**	-0.048**	-0.063**	-0.025**	-0.049**
$ESURP_{t-1}$	-0.737	0.062**	-0.002	-0.008	0.000	-0.010	-0.020**	-0.030**	-0.049**	-0.024**
$EVOL_{t-1}$	-0.665	0.030**	0.008	-0.023**	-0.027**	-0.030**	-0.051**	-0.037**	-0.060**	-0.014*
$NUMEST_{t-1}$	-0.841	0.570**	0.210**	0.082**	0.008	-0.086**	-0.018*	-0.045**	-0.041**	0.088**
$FDISP_{t-1}$	-0.877	0.542**	0.036**	0.044**	0.037**	0.024**	0.005	0.051**	0.052**	0.012
STD_DEV_{t-1}	-0.930	0.012*	0.028**	0.015**	0.034**	0.015**	-0.008	-0.009	-0.004	0.034**

Average log turnover measured over univariate portfolio decile sorts of several controls and disagreement measure. At each month, the cross-section of stocks is assigned to 10 portfolios based on the sorting variable. This procedure is repeated for each month. D_i is the i^{th} decile, $D_j - D_i$ is the difference of average L_TURN in D_j and D_i. Corresponding signficance levels are from a t-test of sample means across corresponding decile pairs. Statistical significance of 5% and 1% are indicated by * and ** respectively

Table 12: **Bivariate Sorts**

	L_TURN_t								
	Control_Var_Ter$_{(1,.)}$			**Control_Var_Ter$_{(2,.)}$**			**Control_Var_Ter$_{(3,.)}$**		
	T_{11}	$T_{12}-T_{11}$	$T_{13}-T_{12}$	T_{21}	$T_{22}-T_{21}$	$T_{23}-T_{22}$	T_{31}	$T_{32}-T_{31}$	$T_{33}-T_{32}$
RET^{+}_{t-1}	-0.978	0.044**	0.006	-0.867	0.060**	0.041**	-0.614	0.061**	-0.019*
RET^{-}_{t-1}	-0.743	-0.046**	-0.062**	-0.958	0.019**	-0.020**	-0.918	0.079**	0.048**
LEV_{t-1}	-1.027	0.113**	0.115**	-0.880	0.078**	0.021**	-0.909	0.116**	-0.046**
$CAPM\_\beta_{t-1}$	-1.137	-0.026**	-0.043**	-0.778	0.010*	-0.072**	-0.458	0.029**	-0.039**
BTM_{t-1}	-0.684	-0.011*	0.014**	-0.945	0.042**	0.053**	-0.974	0.009	-0.007
L_PRC_{t-1}	-1.289	0.082**	0.106**	-0.965	0.178**	0.295**	-0.406	0.125**	0.158**
L_FAGE_{t-1}	-0.721	0.089**	0.015**	-0.870	0.108**	0.021**	-1.156	0.120**	0.043**
$ESURP_{t-1}$	-0.760	0.052**	0.104**	-0.755	0.062**	0.074**	-0.736	-0.057**	-0.028**
$EVOL_{t-1}$	-0.731	0.067**	0.175**	-0.756	0.051**	0.105**	-0.792	-0.058**	-0.024**
$NUMEST_{t-1}$	-0.637	0.126**	0.138**	-0.109	0.122**	0.223**	-0.280	0.195**	0.199**
$FDISP_{t-1}$	-0.667	0.043**	0.022**	-0.325	0.150**	0.194**	-0.228	0.166**	0.140**

Average log turnover measured over bivariate 3×3 sorts. The first sorting variable is one of the 11 control variables and the second sorting variable is the disagreement measure. Sorting is done at time $t-1$ and turnover is measured at time t. The bivariate sorts are dependent sorts where first, one of the control variable is used to make portfolios and then within those portfolios, disagreement is used to make sub-portfolios. Both dimensions of sorting are cross-sectional where a fresh sorting is performed each month. 70/30 portfolio breakpoints are used to make terciles. $T_{i,j}$ is the avergae log turnover for i^{th} control variable tercile and j^{th} disagreement tercile. Each row represents a new control variable being used to make terciles. Within a row, each control variable tercile is represented as a group of three values. There are three groups each corresponding to a different control variable tercile. The second and thrid entries in each group presents difference in average log turnover over two successive terciles. $T_{i,j} - T_{i,j'}$ is the difference between disagreement terciles j and j' within control variable tercile i. Corresponding signficance levels are from a t-test of sample means across corresponding tercile pairs. Statistical significance of 5% and 1% are indicated by * and ** respectively

Table 13: **Variable Definitions**

Variable	Definition
RET^{+} and **RET**$^{-}$	Monthly return is decomposed into two variables based on its sign. $RET^{+} = \max(ret, 0)$ and $RET^{-} = \min(ret, 0)$. ret is adjusted for delisting of firms.
RET_Rf	Excess monthly return. Defined as $RET\_Rf = ret - R_f$ where R_f is the risk-free rate.
DIV_12M	Return accruing to dividend payments in last 12 months. It is calculated as the ratio of past twelve months return by past twelve months price change. The latter doesn't incorporate dividend returns.
PRC	Stock price adjusted for splits, rights issues and other corporate events that affect the face value of a share.
TURN	Monthly share turnover calculated as monthly share volume divided by adjusted shares outstanding.
TURN_1d	Share turnover for the first day of month.
TURN_5d	Share turnover for the first five days (roughly a week) of month.
ILLIQ	Amihud (2002) Illiquidity measure constructed using daily returns and volume data from CRSP daily stock files. $ILLIQ = \frac{1}{D} \cdot \sum_{d=1}^{d=D} \frac{\lvert ret_d \rvert}{vol_d}$, where D is the number of trading days in a month. Days with zero trading are excluded from the summation.
RET_VOL	Monthly return volatility computed as the standard deviation of daily stocks returns.
NUMEST	Number of analyst following a firm in a given month
FDISP	Standard deviation of analyst forecasts following a firm scaled by absolute value of mean forecast estimate. I require that at least two analysts are following the firm ($NUMEST \geq 1$)
RES_COV	Residuals from regressing number of analyst following a stock on firm size and book to market ratio. $\log(1 + NUMEST_t) = \beta_0 + \beta_1 L\_ME_t + \beta_2 BTM_t + u_t$

Table 13: **Variable Definitions** *(continued)*

Variable	Definition
ESURP	Absolute earning surprise is the absolute difference between the most recent quarterly earnings per share (EPS_q) and EPS 4 quarters ago (EPS_{q-4}) scaled by quarter end stock price (P_q). EPS and stock price are adjusted for splits. $ESURP = \frac{\lvert EPS_q - EPS_{q-4}\rvert}{P_q}$ for quarter q.
EVOL	Volatility of earnings is the standard deviation of eight recent quarterly earnings per share scaled by the quarter end stock price. $EVOL = \frac{1}{7 \cdot P_q} \cdot \sum_{i=0}^{7}(EPS_{q-i} - \overline{EPS_q})^2$, where $\overline{EPS_q}$ is the mean EPS over the same period.
L_FAGE	Firm age is the natural log of number of months since the firm first appeared on the CRSP monthly database.
BE and **ME**	Book value and market value of equity construction is described in Appendix A.1.
TURN_GRT	Turnover adjusted as per Gallant et al. (1992). Non-stationarity and calendar effects are removed from both the mean and variance of turnover time-series.[a]
NBUY	Number of anomalies signalling buy (+1) for a firm in a particular month.
NSELL	Number of anomalies signalling sell (-1) for a firm in a particular month.
MEAN_SIGNAL	Average of all signals for a firm in a month. A positive (negative) value indicates that $NBUY(NSELL)$ is higher than $NSELL(NBUY)$.
STD_DEV	Standard deviation of all signals for a firm in a month. I require that at least 10 signals are present to reliably estimate standard deviation.
STD_DEV_Dec	Monthly cross-sectional deciles of STD_DEV.
ABS_DEV	Mean absolute deviation of all the signals for a firm in a month. If T_k represents $k^{th}(1 \le k \le K)$ trading signal then $ABS\_DEV = \frac{1}{K}\sum_{k=1}^{K}\lvert T_k - MEAN\_SIGNAL\rvert$.
CAPM_ALPHA and **CAPM_BETA**	The intercept and slope parameters from regressing firm's excess returns on market excess returns. Regression parameters are obtained in a rolling fashion using the past 60 months of returns data (from t to $t-59$). The intercept is $CAPM\_ALPHA$ while the slope is $CAPM\_BETA$.

Table 13: **Variable Definitions** *(continued)*

Variable	Definition
FF3_ALPHA and **FF3_BETA**	Excess returns are regressed on three well known factors (Fama and French (1992)) to predict returns: excess market return, size and book to market[b]. 60 month rolling regressions are used to estimate the coefficients. $FF3\_ALPHA$ is the intercept while $FF3\_BETA$ is the slope parameter for excess market return.
FF5_ALPHA and **FF5_BETA**	Same procedure to estimate as that of FF-3 regression except that there are two additional factors[c] (Fama and French (2015)).
LO_CORR_SD and **HI_CORR_SD**	Standard deviation of two sets of signals where the two sets differ in their aggregate absolute correlation within the group[d]. At each month, using a brute-force approach over the full set of 36 anomaly signals, I search for a set of 18 signals which produce the smallest aggregate correlation. The standard deviation of signals within this set is LO_CORR_SD and of the remaining set is HI_CORR_SD
PC_DEV	Standard deviation of all signals projected in the principal directions[e].
LO_PCA_SD and **HI_PCA_SD**	Standard deviation of two sets of signals where the partitioning of signals is based on their loadings on the first principal component (PC) of all 36 signals. Deviation of signals in the first half makes LO_PCA_SD while that of the other half makes HI_PCA_SD. The first half has the smallest 18 loadings on first PC and the second half the largest 18.
NUM_FLIPS	The number of times any two pairs of signals flip. Let s and s' be any two distinct signals for a given firm. A flip occurs at time t if $s_{t-1} \cdot s'_{t-1} = -1, s_{t-1} \cdot s_t = -1$ and $s'_{t-1} \cdot s'_t = -1$. The number of such occurrences is NUM_FLIPS.
NUM_DIV	The number of times any two pairs of signals diverge. Let s and s' be any two distinct signals for a given firm. A divergence occurs at time t if $s_{t-1} = s'_{t-1} = 0$ and $s_t \cdot s'_t = -1$. The number of such occurrences is NUM_DIV.
L_TURN_ILLIQ	Residuals from regressing L_TURN on an intercept and L_ILLIQ.
VW_L_TURN	Residuals from regressing L_TURN on an intercept and log of value weighted market turnover[f].
EW_L_TURN	Residuals from regressing L_TURN on an intercept and log of equal weighted market turnover[g].

Table 13: **Variable Definitions** *(continued)*

Variable	Definition
RES_STD_DEV²	Residuals from regressing STD_DEV on an intercept and STD_DEV^2
STD_TURN	Standard deviation of $TURN$ for the last 36 months (from time t to $t-35$)
RES_STD_TURN	Residuals from regressing STD_TURN on an intercept and L_TURN

[a] GRT adjustments is carried out in two steps. In the first step the variable to be adjusted, X is regressed on linear and quadratic time trends as well as calendar month dummies: $X_t \sim \beta_0 + \beta_1 \cdot t + \beta_2 \cdot t^2 + \gamma \cdot D_{1\ldots11} + \epsilon_t$. Here $D_{1\ldots11}$ represents 11 month of the year dummies. In the next step squared residuals are regressed on the same set of variables: $\log(\epsilon_t^2) \sim \beta_0 + \beta_1 \cdot t + \beta_2 \cdot t^2 + \gamma \cdot D_{1\ldots11} + u_t$. Then the GRT adjusted series is defined as, $X\_GRT_t = \exp(u_t/2)$. Finally, X_GRT is linearly transformed so that its mean and variance matches that of X

[b] The explanatory variables of the regression are obtained from Kenneth French's online data library. The SMB (size factor) and HML (value factor) construction is provided at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html

[c] The two additional factors are profitability (RMW) and investments (CMA). Their definition is provided at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5_factors_2x3.html

[d] Let C be the $N \times N$ correlation matrix of the entire set of signals, N being the number of signals. Finding a smaller set of $n(n < N)$ signals is identical to solving the integer program, $\min_v v^T|C|v$, s.t. $v^T\mathbf{1} = n, v \in \{0,1\}^N$, where $\mathbf{1}$ is a vector of all ones and $|C|$ is the absolute correlation matrix with $|C|_{i,j} = |C_{i,j}| \forall i,j \in 1 \ldots N$. The above optimization problem has integer valued range of values v can take. This type of problem is a hard problem and can only solved by iterating over all possible values of v.

[e] At time t, let S be the $F_t \times N$ matrix of signals where F_t is the number of firms and each firm has N signals. Let W be an $N \times N$ matrix of principal components of S. STD_DEV is the row-wise standard deviation of S and PC_DEV is the row-wise standard deviation of SW which is the projection of signals on the principal component space.

[f] Value weighted market turnover is $\sum_{i=1}^{D_t} \frac{ME_{i,t}}{\overline{ME_t}} \cdot TURN_{i,t}$, where $\overline{ME_t} = \sum_{i=1}^{D_t} ME_{i,t}$ and D_t is number of firms at time t.

[g] Equal weighted market turnover at time t is $\frac{1}{D_t} \cdot \sum_{i=1}^{D_t} TURN_{i,t}$, where D_t is number of firms at time t.

Table 14: **Monthly cross-sectional regression: BTM terciles**

	L_TURN_t					
	BTM-1		*BTM-2*		*BTM-3*	
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.449***	1.402***	1.453***	1.359***	1.101***	0.995***
	(0.087)	(0.086)	(0.088)	(0.084)	(0.137)	(0.126)
RET^{-}_{t-1}	-1.806***	-1.729***	-1.797***	-1.658***	-1.688***	-1.587***
	(0.141)	(0.138)	(0.128)	(0.122)	(0.126)	(0.119)
LEV_{t-1}	13.376	16.797	95.487***	84.645***	19.963**	15.309*
	(13.423)	(13.513)	(22.695)	(22.075)	(8.473)	(7.947)
$CAPM\_\beta_{t-1}$	0.333***	0.327***	0.390***	0.373***	0.452***	0.429***
	(0.023)	(0.023)	(0.020)	(0.020)	(0.028)	(0.028)
BTM_{t-1}	39.067	88.277*	-92.039**	-45.653	49.383***	48.728***
	(46.991)	(49.324)	(44.761)	(45.160)	(17.227)	(16.195)
L_PRC_{t-1}	0.169***	0.175***	0.180***	0.206***	0.226***	0.271***
	(0.021)	(0.021)	(0.019)	(0.019)	(0.028)	(0.029)
L_FAGE_{t-1}	-0.076***	-0.067***	-0.008	0.001	-0.025	-0.022
	(0.023)	(0.023)	(0.018)	(0.018)	(0.030)	(0.029)
$ESURP_{t-1}$	0.248***	0.220***	0.521***	0.457***	0.132**	0.123**
	(0.083)	(0.082)	(0.104)	(0.095)	(0.056)	(0.051)
$EVOL_{t-1}$	0.233**	0.237**	0.156	-0.036	0.096	0.054
	(0.110)	(0.109)	(0.165)	(0.148)	(0.071)	(0.065)
$NUMEST_{t-1}$	0.012***	0.013***	0.030***	0.031***	0.042***	0.043***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
$FDISP_{t-1}$	0.119***	0.112***	0.102***	0.089***	0.091***	0.078***
	(0.022)	(0.022)	(0.014)	(0.013)	(0.012)	(0.012)
STD_DEV_{t-1}		0.579***		0.859***		1.310***
		(0.114)		(0.076)		(0.134)
Within R^2	0.122	0.125	0.213	0.222	0.265	0.277
Adj. R^2	0.549	0.550	0.563	0.567	0.550	0.558
$\%R^2$ Explained		0.24		0.74		1.45
Observations	117,365	117,365	178,577	178,577	103,534	103,534

Log turnover regressed on set of controls and several measures of disagreement across three BTM terciles made on 70/30 BTM splits. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 15: **Monthly cross-sectional regression: LEV terciles**

	L_TURN_t					
	LEV-1		*LEV-2*		*LEV-3*	
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.561***	1.461***	1.467***	1.364***	1.053***	0.966***
	(0.087)	(0.082)	(0.091)	(0.086)	(0.133)	(0.124)
RET^{-}_{t-1}	-2.137***	-1.995***	-1.772***	-1.637***	-1.559***	-1.451***
	(0.135)	(0.129)	(0.133)	(0.127)	(0.124)	(0.117)
LEV_{t-1}	837.533***	1204.839***	346.188***	328.359***	11.638*	7.641
	(228.722)	(236.729)	(49.673)	(48.681)	(6.645)	(6.393)
$CAPM\_\beta_{t-1}$	0.410***	0.395***	0.366***	0.350***	0.355***	0.333***
	(0.027)	(0.026)	(0.022)	(0.022)	(0.026)	(0.026)
BTM_{t-1}	-108.482***	-69.161*	-108.916***	-94.678***	29.903**	34.441***
	(37.405)	(37.537)	(23.399)	(22.813)	(12.515)	(12.330)
L_PRC_{t-1}	0.207***	0.217***	0.179***	0.200***	0.222***	0.266***
	(0.023)	(0.023)	(0.020)	(0.020)	(0.026)	(0.027)
L_FAGE_{t-1}	-0.160***	-0.144***	-0.001	0.010	0.070***	0.070***
	(0.025)	(0.025)	(0.020)	(0.020)	(0.027)	(0.027)
$ESURP_{t-1}$	0.369***	0.295***	0.507***	0.465***	0.132***	0.114**
	(0.117)	(0.110)	(0.088)	(0.079)	(0.047)	(0.044)
$EVOL_{t-1}$	0.386***	0.283**	0.181	-0.066	0.182***	0.143**
	(0.149)	(0.144)	(0.171)	(0.161)	(0.064)	(0.060)
$NUMEST_{t-1}$	0.024***	0.023***	0.024***	0.024***	0.034***	0.034***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
$FDISP_{t-1}$	0.128***	0.119***	0.101***	0.086***	0.074***	0.064***
	(0.022)	(0.022)	(0.014)	(0.014)	(0.012)	(0.012)
STD_DEV_{t-1}		0.883***		0.891***		1.286***
		(0.118)		(0.079)		(0.121)
Within R^2	0.224	0.231	0.180	0.191	0.223	0.234
Adj. R^2	0.544	0.548	0.562	0.567	0.555	0.561
$\% R^2$ Explained		0.10		0.32		1.96
Observations	143,316	143,316	162,949	162,949	93,211	93,211

Log turnover regressed on set of controls and several measures of disagreement across three BTM terciles made on 70/30 LEV splits. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 16: **Monthly cross-sectional regression: Analyst following terciles**

	L_TURN_t					
	NUMEST-1		*NUMEST-2*		*NUMEST-3*	
	(1)	(2)	(3)	(4)	(5)	(6)
RET^+_{t-1}	1.506***	1.402***	1.172***	1.093***	1.281***	1.199***
	(0.107)	(0.101)	(0.098)	(0.092)	(0.123)	(0.120)
RET^-_{t-1}	-1.739***	-1.607***	-1.664***	-1.546***	-1.825***	-1.703***
	(0.119)	(0.112)	(0.131)	(0.125)	(0.174)	(0.169)
LEV_{t-1}	34.754***	28.911***	53.392***	45.042***	90.998**	83.743**
	(7.921)	(7.489)	(13.323)	(12.909)	(38.786)	(37.253)
$CAPM\_\beta_{t-1}$	0.390***	0.372***	0.251***	0.235***	0.253***	0.231***
	(0.019)	(0.018)	(0.022)	(0.022)	(0.064)	(0.063)
BTM_{t-1}	5.729	13.498	99.414***	119.245***	98.041	127.337**
	(12.094)	(11.838)	(25.079)	(24.569)	(61.232)	(60.538)
L_PRC_{t-1}	0.188***	0.214***	-0.022	-0.020	-0.052	-0.053
	(0.017)	(0.017)	(0.020)	(0.019)	(0.047)	(0.046)
L_FAGE_{t-1}	-0.001	0.011	-0.083***	-0.071***	-0.206***	-0.190***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.057)	(0.056)
$ESURP_{t-1}$	0.180***	0.147***	0.147	0.111	0.673***	0.624***
	(0.055)	(0.049)	(0.118)	(0.113)	(0.204)	(0.201)
$EVOL_{t-1}$	0.173**	0.124**	0.388*	0.209	0.066	-0.093
	(0.067)	(0.060)	(0.217)	(0.199)	(0.882)	(0.854)
$NUMEST_{t-1}$	0.073***	0.074***	0.003	0.003	-0.011*	-0.012**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.006)	(0.006)
$FDISP_{t-1}$	0.084***	0.071***	0.033**	0.022	0.020	0.014
	(0.011)	(0.010)	(0.015)	(0.015)	(0.025)	(0.024)
STD_DEV_{t-1}		0.996***		0.667***		0.619***
		(0.077)		(0.085)		(0.198)
Within R^2	0.215	0.225	0.142	0.152	0.165	0.173
Adj. R^2	0.511	0.518	0.667	0.670	0.729	0.732
% R^2 Explained		1.08		0.27		-0.01
Observations	268,025	268,025	114,955	114,955	16,496	16,496

Log turnover regressed on set of controls and several measures of disagreement across three 70/30 terciles made on analyst following (*NUMEST*). All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 17: **Monthly cross-sectional regression: Firm Age terciles**

	L_TURN_t					
	FAGE-1		*FAGE-2*		*FAGE-3*	
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.427***	1.291***	1.382***	1.306***	1.567***	1.498***
	(0.088)	(0.081)	(0.119)	(0.114)	(0.095)	(0.093)
RET^{-}_{t-1}	-2.146***	-1.968***	-1.846***	-1.737***	-1.609***	-1.530***
	(0.100)	(0.092)	(0.131)	(0.127)	(0.148)	(0.143)
LEV_{t-1}	15.560*	9.354	34.717***	30.575***	54.536***	49.509***
	(8.025)	(7.833)	(10.420)	(9.993)	(16.919)	(16.343)
$CAPM\_\beta_{t-1}$	0.242***	0.230***	0.416***	0.402***	0.455***	0.443***
	(0.021)	(0.021)	(0.025)	(0.025)	(0.027)	(0.027)
BTM_{t-1}	10.406	23.087	3.842	10.239	4.624	11.789
	(14.320)	(14.219)	(15.586)	(15.437)	(20.579)	(20.569)
L_PRC_{t-1}	0.256***	0.287***	0.197***	0.214***	0.114***	0.127***
	(0.018)	(0.019)	(0.023)	(0.023)	(0.031)	(0.031)
L_FAGE_{t-1}	-0.015	-0.010	0.001	0.006	0.003	0.018
	(0.017)	(0.016)	(0.045)	(0.045)	(0.118)	(0.118)
$ESURP_{t-1}$	0.235***	0.191***	0.202***	0.177***	0.231**	0.178*
	(0.062)	(0.057)	(0.063)	(0.059)	(0.107)	(0.103)
$EVOL_{t-1}$	0.364***	0.278***	0.054	0.007	-0.059	-0.100
	(0.097)	(0.088)	(0.083)	(0.079)	(0.117)	(0.112)
$NUMEST_{t-1}$	0.036***	0.035***	0.029***	0.029***	0.025***	0.025***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$FDISP_{t-1}$	0.104***	0.096***	0.129***	0.120***	0.088***	0.079***
	(0.014)	(0.014)	(0.016)	(0.016)	(0.015)	(0.015)
STD_DEV_{t-1}		1.256***		0.791***		0.538***
		(0.111)		(0.106)		(0.098)
Within R^2	0.218	0.233	0.220	0.227	0.170	0.174
Adj. R^2	0.496	0.505	0.557	0.561	0.535	0.537
% R^2 Explained		1.82		0.85		0.52
Observations	120,166	120,166	149,293	149,293	194,564	194,564

Log turnover regressed on set of controls and several measures of disagreement across three 70/30 terciles made on firm age ($FAGE$). All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 18: **Monthly cross-sectional regression: Report length terciles**

	L_TURN_t					
	LENGTH-1		*LENGTH-2*		*LENGTH-3*	
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.283***	1.191***	0.868***	0.768***	1.156***	1.045***
	(0.191)	(0.184)	(0.190)	(0.173)	(0.153)	(0.138)
RET^{-}_{t-1}	-2.404***	-2.306***	-1.945***	-1.826***	-1.994***	-1.865***
	(0.205)	(0.202)	(0.224)	(0.210)	(0.187)	(0.175)
LEV_{t-1}	28.207*	25.056	22.500**	16.171	2.852	2.514
	(15.496)	(15.546)	(10.102)	(10.055)	(3.298)	(2.916)
$CAPM\_\beta_{t-1}$	0.411***	0.403***	0.329***	0.313***	0.311***	0.287***
	(0.028)	(0.028)	(0.023)	(0.023)	(0.023)	(0.022)
BTM_{t-1}	-56.231**	-48.407*	10.353	18.106	31.550**	33.837**
	(26.207)	(26.141)	(17.321)	(17.253)	(14.074)	(13.859)
L_PRC_{t-1}	0.255***	0.270***	0.211***	0.236***	0.137***	0.163***
	(0.027)	(0.028)	(0.025)	(0.025)	(0.023)	(0.023)
L_FAGE_{t-1}	-0.104***	-0.091***	-0.077***	-0.065***	-0.053**	-0.041**
	(0.027)	(0.027)	(0.022)	(0.022)	(0.021)	(0.020)
$ESURP_{t-1}$	0.037	0.041	0.060	0.035	0.039	0.012
	(0.108)	(0.097)	(0.059)	(0.051)	(0.059)	(0.057)
$EVOL_{t-1}$	0.268*	0.188	0.112*	0.081	0.269***	0.230***
	(0.155)	(0.147)	(0.058)	(0.064)	(0.060)	(0.074)
$NUMEST_{t-1}$	0.032***	0.031***	0.023***	0.022***	0.017***	0.017***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
$FDISP_{t-1}$	0.119***	0.110***	0.126***	0.115***	0.112***	0.098***
	(0.024)	(0.024)	(0.017)	(0.017)	(0.025)	(0.024)
STD_DEV_{t-1}		0.699***		0.964***		1.077***
		(0.122)		(0.109)		(0.096)
Within R^2	0.253	0.258	0.180	0.192	0.155	0.172
Adj. R^2	0.508	0.512	0.467	0.475	0.477	0.488
% R^2 Explained		1.51		2.04		0.36
Observations	23,759	23,759	31,223	31,223	22,527	22,527

Log turnover regressed on set of controls and several measures of disagreement across three 70/30 terciles made on length of quarterly (10-Q) and annual (10-K) reports. Length of the report is measured by number of words in the report. Report length terciles are made every month. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 19: **Monthly cross-sectional regression: EDGAR implementation**

	L_TURN_t			
	pre EDGAR		*post EDGAR*	
	(1)	(2)	(3)	(4)
RET^{+}_{t-1}	1.964***	1.927***	1.117***	1.021***
	(0.111)	(0.111)	(0.104)	(0.095)
RET^{-}_{t-1}	-0.926***	-0.877***	-2.093***	-1.973***
	(0.191)	(0.186)	(0.121)	(0.114)
LEV_{t-1}	51.421***	47.563***	29.602***	25.293***
	(11.876)	(11.825)	(8.587)	(8.308)
$CAPM\_\beta_{t-1}$	0.547***	0.537***	0.372***	0.358***
	(0.034)	(0.033)	(0.019)	(0.019)
BTM_{t-1}	57.493***	62.908***	-1.545	5.043
	(21.200)	(21.047)	(12.986)	(12.903)
L_PRC_{t-1}	0.102***	0.109***	0.214***	0.234***
	(0.029)	(0.029)	(0.018)	(0.019)
L_FAGE_{t-1}	0.137***	0.145***	-0.072***	-0.061***
	(0.031)	(0.031)	(0.018)	(0.018)
$ESURP_{t-1}$	0.290***	0.254***	0.179***	0.153***
	(0.089)	(0.088)	(0.050)	(0.046)
$EVOL_{t-1}$	0.007	-0.047	0.136**	0.091
	(0.144)	(0.145)	(0.060)	(0.056)
$NUMEST_{t-1}$	0.024***	0.024***	0.029***	0.028***
	(0.002)	(0.002)	(0.002)	(0.002)
$FDISP_{t-1}$	0.050***	0.047***	0.139***	0.128***
	(0.012)	(0.012)	(0.015)	(0.014)
STD_DEV_{t-1}		0.373***		0.863***
		(0.094)		(0.083)
Within R^2	0.164	0.166	0.229	0.238
Adj. R^2	0.278	0.280	0.438	0.445
$\%R^2$ Explained		0.34		0.98
Observations	127,868	127,868	243,711	243,711

Log turnover regressed on different set of explanatory variables for two disjoint periods: pre EDGAR from Jan 1976 to Mar 1993 and post EDGAR from Jun 1996 to Dec 2019. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 20: **Monthly cross-sectional regression: Short-sale constraint terciles**

	L_TURN_t					
	$INST\_OWN-1$	$INST\_OWN-2$	$INST\_OWN-3$	$SHORT\_INT-1$	$SHORT\_INT-2$	$SHORT\_INT-3$
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.631***	1.507***	1.243***	1.256***	1.149***	1.221***
	(0.190)	(0.109)	(0.086)	(0.156)	(0.079)	(0.088)
RET^{-}_{t-1}	-1.802***	-1.929***	-1.854***	-1.366***	-1.489***	-1.692***
	(0.162)	(0.145)	(0.152)	(0.161)	(0.115)	(0.141)
LEV_{t-1}	17.015***	17.355*	31.908**	37.407***	9.805	1.576
	(6.469)	(9.851)	(13.573)	(11.584)	(9.136)	(5.770)
$CAPM\_\beta_{t-1}$	0.295***	0.284***	0.207***	0.308***	0.216***	0.156***
	(0.033)	(0.019)	(0.020)	(0.027)	(0.015)	(0.018)
BTM_{t-1}	-1.864	-6.476	22.124	8.663	51.788***	17.519
	(13.789)	(9.907)	(25.859)	(19.422)	(17.704)	(12.276)
L_PRC_{t-1}	0.117***	0.066***	0.052*	0.129***	0.119***	0.098***
	(0.024)	(0.021)	(0.031)	(0.028)	(0.016)	(0.018)
L_FAGE_{t-1}	-0.012	-0.006	-0.031**	-0.007	0.004	-0.003
	(0.025)	(0.012)	(0.015)	(0.019)	(0.010)	(0.011)
$ESURP_{t-1}$	0.218***	0.271***	0.197***	0.081	0.034	0.247***
	(0.084)	(0.082)	(0.072)	(0.117)	(0.088)	(0.060)
$EVOL_{t-1}$	0.342***	-0.163**	0.032	0.180	0.014	-0.054
	(0.101)	(0.075)	(0.115)	(0.179)	(0.083)	(0.061)
$NUMEST_{t-1}$	0.035***	0.018***	0.020***	0.039***	0.018***	0.019***
	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
$FDISP_{t-1}$	0.088***	0.057***	0.031***	0.074***	0.010	0.021**
	(0.021)	(0.012)	(0.012)	(0.019)	(0.012)	(0.010)
STD_DEV_{t-1}	1.493***	0.961***	0.860***	0.093	0.488***	0.995***
	(0.177)	(0.086)	(0.087)	(0.113)	(0.071)	(0.084)
Within R^2	0.163	0.146	0.136	0.167	0.121	0.131
Adj. R^2	0.379	0.533	0.594	0.420	0.520	0.558
% R^2 Explained	1.91	1.65	1.79	0.01	0.55	2.00
Observations	21,253	60,129	49,725	30,841	59,051	41,215

Log turnover regressed on set of controls and several measures of disagreement across short sale constraint terciles. First three specifications are across 13-F institutional ownership terciles while the last three specifications are across short interest terciles. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 21: **Monthly cross-sectional regression: Time since 10-K filing**

	L_TURN_t			
	$s \in \{1,2,3\}$	$s \in \{4,5,6\}$	$s \in \{7,8,9\}$	$s \in \{10,11,0\}$
	(1)	(2)	(3)	(4)
RET^{+}_{t-1}	0.877***	1.161***	1.010***	1.091***
	(0.199)	(0.119)	(0.112)	(0.116)
RET^{-}_{t-1}	-1.908***	-2.107***	-1.856***	-2.027***
	(0.146)	(0.142)	(0.243)	(0.138)
LEV_{t-1}	38.664***	39.273***	36.573***	18.328**
	(12.655)	(10.938)	(9.905)	(8.330)
$CAPM\_\beta_{t-1}$	0.361***	0.359***	0.344***	0.351***
	(0.020)	(0.020)	(0.020)	(0.021)
BTM_{t-1}	-5.355	9.156	12.004	1.860
	(13.615)	(18.660)	(17.297)	(12.903)
L_PRC_{t-1}	0.240***	0.256***	0.238***	0.219***
	(0.021)	(0.023)	(0.022)	(0.023)
L_FAGE_{t-1}	-0.072***	-0.071***	-0.073***	-0.063***
	(0.019)	(0.019)	(0.019)	(0.019)
$ESURP_{t-1}$	0.147**	0.125	0.162*	0.095
	(0.065)	(0.089)	(0.088)	(0.081)
$EVOL_{t-1}$	-0.063	0.056	0.229***	0.287***
	(0.075)	(0.063)	(0.075)	(0.109)
$NUMEST_{t-1}$	0.027***	0.027***	0.029***	0.028***
	(0.002)	(0.002)	(0.002)	(0.002)
$FDISP_{t-1}$	0.114***	0.122***	0.142***	0.158***
	(0.017)	(0.020)	(0.021)	(0.025)
STD_DEV_{t-1}	0.882***	0.881***	0.928***	0.768***
	(0.089)	(0.088)	(0.090)	(0.090)
Within R^2	0.239	0.245	0.237	0.225
Adj. R^2	0.464	0.460	0.451	0.438
$\% R^2$ Explained	1.23	1.20	1.32	0.93
Observations	58,405	57,683	57,438	47,828

Log turnover regressed controls and disagreement across four samples based on time since the last 10-K filing. s is the number of months since the last filing of form 10-K. $s = 0$ represents the month of filing. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 22: **Monthly cross-sectional regression: NASDAQ Stocks**

	L_TURN_t					
	(1)	(2)	(3)	(4)	(5)	(6)
RET^{+}_{t-1}	1.451***	1.500***	1.332***	1.359***	1.388***	1.231***
	(0.070)	(0.069)	(0.067)	(0.066)	(0.066)	(0.063)
RET^{-}_{t-1}	-2.017***	-2.023***	-2.198***	-1.876***	-1.876***	-2.042***
	(0.081)	(0.083)	(0.081)	(0.078)	(0.079)	(0.078)
LEV_{t-1}	37.170***	44.810***	41.451***	29.127***	33.748***	32.553***
	(10.359)	(9.834)	(11.121)	(9.973)	(9.474)	(10.703)
$CAPM\_\beta_{t-1}$	0.324***	0.328***	0.410***	0.314***	0.316***	0.398***
	(0.012)	(0.012)	(0.014)	(0.012)	(0.012)	(0.014)
BTM_{t-1}	-100.411***	-84.542***	-101.308***	-79.525***	-67.652***	-78.373***
	(15.120)	(14.414)	(16.782)	(14.655)	(13.855)	(16.269)
L_PRC_{t-1}	0.204***	0.195***	0.374***	0.222***	0.217***	0.393***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)
L_FAGE_{t-1}	-0.129***	-0.129***	-0.124***	-0.113***	-0.113***	-0.107***
	(0.019)	(0.019)	(0.022)	(0.019)	(0.019)	(0.022)
$ESURP_{t-1}$	0.307***		0.368***	0.281***		0.338***
	(0.058)		(0.061)	(0.053)		(0.056)
$EVOL_{t-1}$	0.181**		0.232***	0.055		0.094
	(0.078)		(0.082)	(0.078)		(0.082)
$NUMEST_{t-1}$	0.055***	0.056***		0.055***	0.055***	
	(0.003)	(0.003)		(0.003)	(0.003)	
$FDISP_{t-1}$	0.098***	0.100***		0.094***	0.095***	
	(0.009)	(0.009)		(0.009)	(0.009)	
STD_DEV_{t-1}				1.060***	1.102***	1.167***
				(0.077)	(0.076)	(0.082)
Within R^2	0.286	0.284	0.208	0.293	0.293	0.217
Adj. R^2	0.453	0.452	0.393	0.459	0.458	0.400
$\% R^2$ Explained				3.54	3.73	4.00
Observations	378,643	378,643	378,643	378,643	378,643	378,643

Log turnover regressed on different set of explanatory variables on NASDAQ stocks. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. % R^2 explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 23: **Monthly cross-sectional regression: disagreement by different stock splits**

	L_TURN_t						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$^{ALL}STD\_DEV^{80/20}_{t-1}$		0.854***					
		(0.066)					
$^{ALL}STD\_DEV^{70/30}_{t-1}$			0.866***				
			(0.073)				
$^{ALL}STD\_DEV^{50/50}_{t-1}$				-0.038			
				(0.097)			
$^{NYSE}STD\_DEV^{80/20}_{t-1}$					0.714***		
					(0.063)		
$^{NYSE}STD\_DEV^{70/30}_{t-1}$						0.755***	
						(0.072)	
$^{NYSE}STD\_DEV^{50/50}_{t-1}$							-0.292***
							(0.096)
Within R^2	0.198	0.210	0.207	0.198	0.208	0.205	0.198
Adj. R^2	0.535	0.542	0.540	0.535	0.541	0.539	0.535
$\% R^2$ Explained		2.10	1.59	-0.00	1.00	0.80	0.03
Observations	399,476	399,476	399,476	399,476	399,476	399,476	399,476

Log turnover regressed on set of controls and disagreement constructed using either NYSE/AMEX or ALL (including NASDAQ as well) stock universe as well as different stock split criterion. A 70/30 stock split assigns top 30% stocks to buy category, bottom 30% stocks to sell category and the rest to hold category. Similarly for 80/20 and 50/50 splits. All independent variables are one month lagged variables. Definitions of all the variables appears in Appendix A.2. All regression specifications have industry and year fixed effects. Within R^2 is the R^2 for within groups variation with fixed effects projected out while Adj. R^2 is for the entire model. $\% R^2$ explained is the percent of unexplained variation explained by including disagreement in the regression. Standard errors are double clustered by firm and year-month. Statistical significance of 10%, 5% and 1% are indicated by *, ** and *** respectively.

Table 24: **Univariate Sorts: Different disagreement measures**

	L_TURN_t									
	D_1	$D_2 - D_1$	$D_3 - D_2$	$D_4 - D_3$	$D_5 - D_4$	$D_6 - D_5$	$D_7 - D_6$	$D_8 - D_7$	$D_9 - D_8$	$D_{10} - D_9$
STD_DEV_{t-1}	-0.930	0.012*	0.028**	0.015**	0.034**	0.015**	-0.008	-0.009	-0.004	0.034**
ABS_DEV_{t-1}	-0.961	0.010	0.051**	0.043**	0.029**	0.011*	0.000	-0.017**	-0.006	0.034**
PC_DEV_{t-1}	-0.829	0.004	0.023**	0.013	0.032**	0.021**	0.030**	-0.011	-0.027**	-0.045**
$LO\_CORR\_SD_{t-1}$	-1.098	0.039**	0.222**	-0.058**	0.055**	-0.033**	0.022**	-0.007	-0.004	-0.030**
$HI\_CORR\_SD_{t-1}$	-1.030	0.053**	0.106**	-0.022**	0.006	0.059**	0.016**	0.008	0.010	-0.114**
NUM_FLIPS_{t-1}	-0.965	0.093**	-0.009	0.020	-0.020	0.001	-0.095*	0.124	-0.134	0.069
NUM_DIV_{t-1}	-1.039	0.131**	0.067**	0.048**	0.036**	-0.000	0.012	0.025	0.061*	0.088**

Average log turnover measured over univariate portfolio decile sorts of several measures of disagreement. At each month, the cross-section of stocks is assigned to 10 portfolios based on the sorting variable. This procedure is repeated for each month. D_i is the i^{th} decile, $D_j - D_i$ is the difference of average L_TURN in D_j and D_i. Corresponding signficance levels are from a t-test of sample means across corresponding decile pairs. Statistical significance of 5% and 1% are indicated by * and ** respectively

Table 25: **Univariate Sorts: Different turnover measures**

	Deciles made on STD_DEV_{t-1}									
	D_1	D_2-D_1	D_3-D_2	D_4-D_3	D_5-D_4	D_6-D_5	D_7-D_6	D_8-D_7	D_9-D_8	$D_{10}-D_9$
$L\_TURN\_1d_t$	0.455	0.021**	0.034**	0.012*	0.031**	0.013*	-0.018**	-0.015*	-0.020**	0.022**
$L\_TURN\_5d_t$	2.170	0.006	0.023**	0.013*	0.028**	0.014*	-0.015**	-0.019**	-0.016*	0.019*
L_TURN_t	-0.930	0.012*	0.028**	0.015**	0.034**	0.015**	-0.008	-0.009	-0.004	0.034**
$L\_TURN\_GRT_t$	-2.727	0.071**	0.030**	0.055**	0.056**	0.042**	0.072**	0.084**	0.079**	0.094**
$L\_TURN\_D_t$	-0.037	0.013**	0.012**	0.011**	0.011**	-0.002	0.004	-0.006	0.005	0.008
$L\_TURN\_ILLIQ_t$	-0.058	0.013**	0.011**	0.008**	0.013**	0.010**	0.014**	0.024**	0.049**	0.052**
$VW\_L\_TURN_t$	-0.040	0.016**	0.007**	0.011**	0.009**	0.003	0.006*	-0.005	0.003	0.007
$EW\_L\_TURN_t$	-0.025	0.008**	0.003	0.009**	0.009**	0.002	0.002	-0.006*	-0.000	0.003

Several measures of turnover averaged over univariate decile sorts of STD_DEV_{t-1}. At each month, the cross-section of stocks is assigned to 10 portfolios based on the disagreement and then several measures of turnover are avergaed over these portfolios. This procedure is repeated for each month. In the first row, D_i is the i^{th} decile, D_j-D_i is the difference of average L_TURN in D_j and D_i. Similarly for other rows as well. Corresponding signficance levels are from a t-test of sample means across corresponding decile pairs. Statistical significance of 5% and 1% are indicated by * and ** respectively